Essays zur Idee der Universität im 21. Jahrhundert

Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz



Herausgegeben von

Gabi Reinmann und Alice Watanabe

Universität Hamburg 2026

Empfohlener Zitierhinweis

Reinmann, G. & Watanabe, A. (2026) (Hrsg.). *Essays zur Idee der Universität im 21. Jahrhundert: Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz*. Universität Hamburg. https://doi.org/10.25592/uhhfdm.22184

Angaben zur Förderung

Wir danken der Claussen-Simon-Stiftung sowie der Hamburgischen Wissenschaftlichen Stiftung herzlich für ihre Unterstützung und Förderung.

Copyright- und Lizenzhinweise

Die Lizenz CC BY-ND erlaubt das Teilen und Verbreiten von Inhalten, auch für kommerzielle Zwecke. Voraussetzung sind die Nennung der Urheber:innen und die unveränderte Weitergabe des Materials. Bearbeitungen oder Veränderungen sind nicht erlaubt.
Weitere Informationen: https://creativecommons.org/licenses/by-nd/4.0/

Inhalt

Einleitung

Gabi Reinmann & Alice Watanabe……3

Auftakt: Annäherungen an Wissenschaft, Bildung und KI

Wissenschaft, Bildung und KI – ein Kaleidoskop
Rüdiger Rhein……7

Von der Schockstarre zur Symbiose?! Acht Gespräche über generative KI in der Hochschule
Andreas Hebbel-Seeger……16

Teil I: Was KI mit Wirklichkeit und Erkenntnis macht

Die maschinelle Konstruktion der Wirklichkeit
Antje Michel……24

Kommentar zum Essay „Die maschinelle Konstruktion der Wirklichkeit“ von Antje Michel
Michael Pandey……30

Generative KI und die Verführung der Reibungslosigkeit
Judith Simon……37

Von der Effizienz-Illusion und Bot-Delusion in Forschung und Lehre
Isabel Steinhardt, Christian Schneijderberg & Oliver Wieczorek……44

Ambivalenzen des technologischen Fortschritts
Christian Leineweber……49

Der ambivalente Blick in den KI-Spiegel – mit Bezugnahme auf Christian Leineweber
Daniel Rütten……56

Veränderte Mensch-Technologie-Relationen – Perspektiven auf multistabile und imitierende Technologien
Silke Elisabeth Wrede……62

Teil II: Wahrheit, Verantwortung, Bildung

Die Universität als „Instanz der Wahrheit schlechthin“ – Wahrheit als zentraler Wert der Universität
Markus Stepanians……69

Wie KI die Universität als „Instanz der Wahrheit“ beeinflusst – ein Kommentar zu Markus Stepanians
Gabi Reinmann & Frank Vohle……75

Wahrheit und der explanatorische Primat der Welt – eine Replik auf den Kommentar von Gabi Reinmann und Frank Vohle
Markus Stepanians……81

Die Zukünfte der Universität: zwischen technisierter und lebendiger Wissenschaft
Dominikus Herzberg……..*86*

Ach du liebe Wahrheit – eine Nachbetrachtung aus dem Dialog mit Markus Stepanians
Dominikus Herzberg……..*92*

Wahrheitsapparate und Effektivität – zu zwei Thesen von Dominikus Herzberg
Markus Stepanians……..*95*

Institutionelle Verantwortung in der Wissenschaft und in der KI-Nutzung – Fragen ans Bildungsverständnis
Ines Langemeyer……..*101*

Der vergessene Faktor Zeit. Der universitäre Paradigmenwechsel von der Bildung zur Ausbildung und die Rolle von KI
Dorothea Winter……..*106*

Teil III: Denkfiguren: Utopie, Analogie, Gedankenexperiment

Die Universität als Soziale Plastik – eine utopische Skizze
Nicolaus Wilder……..*115*

Universitätsidee im Spiegel der Sportidee. Eine Analogie als Inspiration
Frank Vohle……..*122*

Über die Idee der Bildung – eine Nachbetrachtung zum Beitrag „Universitätsidee im Spiegel der Sportidee“ von Frank Vohle
Dominikus Herzberg……..*129*

Die körperlose Universität – ein Gedankenexperiment
Gabi Reinmann & Alice Watanabe……..*133*

Die Idee der Universität Mitte des 21. Jahrhunderts? Kommentar einer Geistes- und Kulturwissenschaftlerin, [hm]niedergeschrieben am 3. Mai 2038
Petra Gehring……..*140*

Ausklang

Das Symposium als Bricolage: Eine zu stärkende Praxis für die Zukunft von Universitäten
Alinde Keller……..*145*

Einleitung

Gabi Reinmann & Alice Watanabe

Dieser Essayband ist auf ungewöhnliche Weise entstanden und unterscheidet sich in vielem von einem klassischen Sammelband. Sein Ausgangspunkt ist ein Symposium zur Idee der Universität in Zeiten Künstlicher Intelligenz, das im Juli 2026 (mit Unterstützung der Claussen-Simon-Stiftung und der Hamburgischen Wissenschaftlichen Stiftung) in Hamburg stattfand und dabei besonders die Hochschulbildung in den Blick nahm. Wir hatten elf Essayistinnen und Essayisten eingeladen, vorab einen Beitrag zu verfassen; hinzu kam ein eigener Text von uns. Diese Essays bildeten die Grundlage für Diskussionen mit rund 45 Wissenschaftlerinnen und Wissenschaftlern aus unterschiedlichen Disziplinen.

Weder mit dem Symposium noch mit dem Band verfolgten wir ein vorgegebenes Ziel. Es ging uns darum, als freie und gleiche Mitglieder von Hochschulen in einen Diskurs über die Universität in Zeiten von KI einzutreten, in dessen Mittelpunkt die Diskussion selbst stand: das Kennenlernen anderer und auch fremder Positionen und der Austausch über Disziplingrenzen hinweg. Das Symposium wie auch den Band verstanden und verstehen wir als Vorhaben, die ihren Sinn in sich selbst tragen.

Eine Veröffentlichung der Essays hatten wir im Vorfeld als Möglichkeit angedacht, ohne es fest geplant zu haben. Offen war entsprechend auch, welche Form der potenzielle Band annehmen würde. Im abschließenden Plenum des Symposiums entschieden wir gemeinsam, die Essays zu publizieren und den Band für weitere Essays ebenso wie für Kommentare zu den bereits vorliegenden Texten zu öffnen – eine Erweiterung, die vorab nicht vorgesehen war; sie ist aus den Begegnungen und Diskussionen auf dem Symposium hervorgegangen. Der Band sollte den Diskurs über die Idee der Universität und die Hochschulbildung in Zeiten von KI fortführen und den Teilnehmenden einen Ort bieten, an dem sie über die Veranstaltung hinaus schriftlich weiterdenken und miteinander im Gespräch bleiben können.

Dieses Vorgehen ähnelt Hannah Arendts Überlegungen zum Handeln. Die politische Denkerin versteht unter Handeln eine Tätigkeit, in der Menschen als Freie und Gleiche miteinander in Beziehung treten, ihre unterschiedlichen Perspektiven auf die Welt teilen und gemeinsam etwas Neues beginnen (Arendt, 2013, S. 42f.; 215f.; 226; 295). Dabei entsteht ein Gewebe menschlicher Beziehungen (Arend, 2013, S. 225) dessen Formen und Muster sich nicht vorhersehen lassen und sich stetig weiterentwickeln. Arendts Überlegungen gelten zwar dem Politischen und nicht der Wissenschaft, bieten aber eine Sprache, in der sich manches an unserem Vorhaben gut beschreiben lässt.

Für den Band hieß das zunächst, allen Teilnehmenden die Möglichkeit zu geben, sich mit neuen Essays oder mit Kommentaren zu den vorliegenden Texten einzubringen. Mehrere sind dieser Einladung gefolgt: Mit den Texten von Andreas Hebbel-Seeger, Silke Elisabeth Wrede und Alinde Keller kamen drei neue Beiträge hinzu, mit den Kommentaren von Michael Pandey und Daniel Rütten *neue* Stimmen mit Bezug zu bereits vorliegenden Texten.

Abgesehen von wenigen formalen Vorgaben, etwa zum Umfang, haben wir bewusst auf ein klassisches Review-Verfahren verzichtet und uns auf eine redaktionelle Durchsicht beschränkt. Damit wollten wir die Autorinnen und Autoren als freie und gleiche Akteure mit ihren Positionen ernst nehmen und sie zugleich an der Verantwortung beteiligen, die damit einhergeht.

Das bedeutete allerdings auch, damit zu leben, dass sich manche unserer Vorstellungen und Wünsche nicht oder nur teilweise verwirklichen ließen, und mit Entwicklungen umzugehen, die sich erst aus dem Prozess ergaben. Auf dem Symposium waren alle Essays auf lebhafte Resonanz gestoßen: Sie wurden intensiv diskutiert und gaben Anlass zu weiteren thematischen Debatten. Die schriftliche Kommentierung dagegen fiel verhaltener aus, als sie hätte sein können; manche Essays wurden ausführlich kommentiert, andere gar nicht. Das mag auch damit zusammenhängen, dass uns für das schriftliche, öffentliche Kommentieren der Texte von Kolleginnen und Kollegen derzeit eine eingeübte Kommunikationskultur fehlt. Wir sind es gewohnt, auf Tagungen mündlich zu diskutieren oder im Peer Review anonym zu begutachten, kaum aber, die Texte anderer öffentlich und unter eigenem Namen zu kommentieren – dies nicht in der Absicht, die Güte eines Textes zu bewerten, sondern die darin formulierten Gedanken aufzunehmen und fortzuführen. Hier haben wir die Unvorhersehbarkeit und Spontanität gemeinsamen Handelns, von der Arendt spricht, unmittelbar erfahren. Zugleich verdanken wir demselben Vorgehen Denkanstöße und Einsichten, mit denen wir nicht gerechnet hatten und die auf anderem Weg wohl nicht entstanden wären.

Am deutlichsten zeigt sich das an den Dialogen, die sich zwischen einzelnen Texten entwickelt haben. Zu mehreren Essays sind Kommentare und Nachbetrachtungen auch unter den Essayisten entstanden, die von Anfang an dabei waren. Am weitesten reichte der Austausch um Markus Stepanians' Essay zur Universität als „Instanz der Wahrheit": Auf den Kommentar von Gabi Reinmann und Frank Vohle zu diesem Beitrag antwortete Stepanians mit einer Replik, und zwischen ihm und Dominikus Herzberg entspann sich ein Wechsel aus Kommentar und Nachbetrachtung. Diese Dialoge verliefen asynchron und schriftlich und erstreckten sich über Wochen. Anders als im Gespräch auf einer Tagung eröffnen solche „langsamen" Dialoge größere Zeiträume: Zeit, den Gedanken des Gegenübers erst einmal nachzuvollziehen, neu entstehende Argumente zu prüfen, sie zu formulieren und mitunter wieder zu verwerfen. Gerade dadurch wurde die Reflexion in besonderer Weise angeregt: bei denen, die geschrieben haben, und, so hoffen wir, auch bei denen, die nun lesen. Wir haben die Kommentare deshalb nicht in einem eigenen Teil gesammelt, sondern direkt hinter die Texte gestellt, auf die sie sich beziehen, damit sich der Austausch beim Lesen auch verstehen lässt. So berühren diese Dialoge eine Frage, die uns schon auf dem Symposium beschäftigt hat: was Denken an der Universität ausmacht, wenn generative KI in Sekunden und scheinbar mühelos Texte hervorbringt. Sie zeigen, was entsteht, wenn wir uns zum Denken und zum Urteilen Zeit nehmen und Reibung zulassen.

Gegliedert ist der Band in drei Teile, die von einem Auftakt und einem Ausklang gerahmt werden.

Im *Auftakt: Annäherungen an Wissenschaft, Bildung und KI* dreht Rüdiger Rhein das Kaleidoskop aus Wissenschaft, Bildung und KI und zeigt, wie viele Muster sich aus diesen drei Elementen ergeben. Andreas Hebbel-Seeger lässt in acht Gesprächen diejenigen zu Wort kommen, die den Umgang mit generativer KI an Hochschulen gerade aushandeln, irgendwo zwischen Schockstarre und Symbiose.

Teil I: Was KI mit Wirklichkeit und Erkenntnis macht umfasst Beiträge von Antje Michel, Michael Pandey, Judith Simon, Isabel Steinhardt et al., Christian Leineweber, Daniel Rütten und Silke Elisabeth Wrede. Zusammen betrachtet fragen sie, wie generative KI unseren Zugang zur Wirklichkeit verändert, wenn Maschinen an deren Konstruktion mitwirken und Erkenntnis plötzlich mühelos zu haben scheint. Dabei richtet sich der Blick auf das Verführerische der Reibungslosigkeit und auf die Illusion der Effizienz, aber auch auf die Ambivalenzen des technologischen Fortschritts und darauf, wie sich die Beziehung zwischen Mensch und Technologie verschiebt.

Die Beiträge in *Teil II: Wahrheit, Verantwortung, Bildung* stammen von Markus Stepanians, Gabi Reinmann und Frank Vohle, Dominikus Herzberg, Ines Langemeyer und Dorothea Winter. Gemeinsam rücken sie die Universität selbst in den Mittelpunkt: als „Instanz der Wahrheit", deren Anspruch sich durch KI neu stellt und um den die Autoren in Kommentar, Replik und Nachbetrachtung miteinander ringen. Von der Wahrheit führt der Weg zur institutionellen Verantwortung und schließlich zur Frage, ob die Universität noch bildet oder nur noch ausbildet, ein Ort der technisierten oder lebendigen Wissenschaft ist und welche Rolle die vergessene Zeit dabei spielt.

Teil III: Denkfiguren: Utopie, Analogie, Gedankenexperiment führt Beiträge von Nicolaus Wilder, Frank Vohle, Dominikus Herzberg, Gabi Reinmann und Alice Watanabe sowie Petra Gehring zusammen. Ihnen ist gemeinsam, dass sie vom Argument zum Entwurf wechseln und die Universität im Modus des Möglichen denken: im Spiegel des Sports, als körperlose Institution, als soziale Plastik. Am Ende steht ein Blick aus dem Jahr 2038, der die Gegenwart des Bandes von ihrer möglichen Zukunft aus betrachtet.

Im *Ausklang* wendet sich der Band seinem eigenen Entstehen zu: Alinde Keller versteht das Symposium als Bricolage, als ein Zusammensetzen aus dem, was da ist. Damit wird die Form, in der dieser Band entstanden ist, selbst zu einem Vorschlag für die Universität der Zukunft.

Literatur

Arendt, H. (2013). *Vita activa oder Vom tätigen Leben* (11. Aufl.). Piper.

Auftakt

Annäherungen an Wissenschaft, Bildung und KI

Wissenschaft, Bildung und KI – ein Kaleidoskop

Rüdiger Rhein

„Die Idee der Universität im 21. Jahrhundert – Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz“. Mit diesem Titel ging das Symposium aufs Ganze, denn hier wurden vier große Themenblöcke kombiniert, die auch jeweils für sich einer eigenen Erörterung würdig sind. Ich versuche im Folgenden, einige Relevanzstrukturen dieser Themenkreise zu benennen und Dimensionen aufzuzeigen, in denen sich ein Diskurs bewegen könnte. Als Hintergrundfrage mag gelten: Welche grundlegenden Veränderungen induziert generative KI in den menschlichen Welt-, Sozial- und Selbstverhältnissen im Allgemeinen? Und was bedeutet dies für akademische Bildung im Besonderen? Dabei ist zu unterscheiden, was genuin neu ist, und was durch generative KI lediglich verschärft, aber nicht originär hervorgerufen wird. Angesichts der Unabgeschlossenheit dieser Entwicklungen soll die Metapher des Kaleidoskops verwendet werden, das je nach Drehwinkel unterschiedliche Ansichten generiert.

1. Die Veränderung des Technotops durch generative KI

Mit dem aktuell erreichten Entwicklungsstand der KI sind datenverarbeitende Maschinen verfügbar, die durch die technische Realisierung maschinellen Lernens auf konnektionistischer Basis Rechenoperationen durchführen können, deren Ergebnisse analog sind zu den Ergebnissen bestimmter menschlicher kognitiver Prozesse: Bild- und Spracherkennung sowie allgemeine Mustererkennung, die Generierung von Texten und sonstigen sprachlichen Äußerungen, Übersetzungen zwischen verschiedenen Sprachen, die Generierung von Bildern, Videos und anderen Artefakten, Informationsrecherche und Darstellung von Rechercheergebnissen u.a.m. Die Resultate dieser Rechenoperationen lassen sich als Bearbeitung kognitiver Aufgaben lesen, die an die KI delegiert worden sind. Unbenommen bleibt die Frage, von welcher Qualität diese Ergebnisse sind; eine anschließende Prüfung, Beurteilung und kritische Reflektion sind empfehlenswert, verbleiben aber im Ermessen der Nutzenden. Mit KI stehen damit technische Assistenzsysteme zur Verfügung, die auf menschlichen Input eine spezifische Resonanz erzeugen. Zwar handelt es sich nur um maschinelle Datenmanagementagenten, aufgrund ihrer funktionellen Leistungen, die u.a. durch maschinelles Lernen ermöglicht werden, haben diese Systeme aber den Charakter von teilautonomen Ko-Akteuren, mit denen eine spezielle Form von Mensch-Maschine-Interaktion möglich ist.

Mit der niederschwelligen Verfügbarkeit und der weitreichenden Durchdringungstiefe des KI-Einsatzes stellen sich Fragen der Technikbewertung: Welches Technotop entsteht durch den umfassenden Einsatz von KI? Welche Ressourcen werden durch generative KI geschaffen, und welche gehen verloren? Welche Risiken birgt diese Technologie und inwiefern sind diese Risiken beherrschbar? Mit welchen unbeabsichtigten Nebenfolgen ist zu rechnen (sofern sich diese derzeit überhaupt absehen lassen)? Welche ethischen Fragen

stellen sich im Hinblick auf den Einsatz von KI? Und welche Auswirkungen hat dieses Technotop auf menschliche Welt-, Sozial- und Selbstbeziehungen?

Bei diesen Fragen ist zu berücksichtigen, dass technologische Entwicklungen immer auch mit Unsicherheiten und Unwägbarkeiten einhergehen, die zugespitzt sowohl in dystopischen Befürchtungen als auch in Heilserwartungen münden können. Gleichwohl gilt am Ende: „Zuerst ist Technik neu, dann gewohnt, dann unsichtbar." (Kornwachs, 2013, S. 2). Neue Technologien werden in einen soziotechnischen Gesamtkomplex inkludiert. Dabei zeigen sich typische Muster: Routinisierung und Veralltäglichung des Technologieeinsatzes aufgrund einschlägiger Vorteile, insbesondere Verfügungserweiterungen, Effizienzsteigerungen und Handlungsentlastung – in einem Spektrum von unhinterfragten technokratischen Adaptionen, pragmatischer Nutzung und informiertem, reflexiv-kritischem Umgang. Möglich sind darüber hinaus Entwicklungen von eigensinnigem Nutzungsverhalten, die aktuell noch nicht absehbar sind. Ebenso können sich dysfunktionale Momente zeigen, die Fragen nach einer Regulierung aufwerfen. Vorstellbar ist außerdem, dass sich gegen dystopische Momente Tendenzen romantischer Rückbesinnung entwickeln. Und nicht zuletzt ist mit Agglomerationstendenzen von Anbietern und ökonomischen Interessen zu rechnen, die Gewinner und Verlierer hervorbringen.

Aktuell konzentriert sich die Diskussion u.a. auf Phänomene der Vermenschlichung generativer KI im Zuge ihrer Nutzung sowie auf die Verführung zur Bequemlichkeit bei der Bearbeitung anspruchsvoller geistiger Aufgaben (skill skipping; Nuxoll, 2024) und dem damit einhergehenden Risiko des Verlustes von Fähigkeiten und Fertigkeiten (deskilling; Reinmann, 2023).

2. KI und Wissenschaft

KI ist ein wissenschaftliches Forschungsthema in der Informatik, in der Philosophie, in den Sozialwissenschaften, in der Sprachwissenschaft, in den Bildungswissenschaften und weiteren Disziplinen. Das Forschungsspektrum umfasst u.a. die Weiterentwicklung der Technologie und ihrer Funktionsprinzipien, die Leistungsfähigkeit maschinellen Lernens und die Qualität der generierten Artefakte, Analysen des „digitalen Wissens" und der „maschinellen Epistemologie" (Gramelsberger & Sigwart, 2025, S. 261 ff.), die Beforschung von Veränderungen der kollektiven Wissenssysteme und der generellen Auswirkungen des Einsatzes von KI in den unterschiedlichen gesellschaftlichen Teilbereichen unter technologischen, ethischen, rechtlichen, ökonomischen und weiteren Perspektiven. Hinzu kommt die wissenschaftliche Selbstreflexion im Hinblick auf die Frage, wie sich wissenschaftliches Wissen selbst durch KI verändert.

Dies verweist auf die Verwendung von KI in den einzelnen wissenschaftlichen Disziplinen zum Zwecke des wissenschaftlichen Arbeitens auf der Grundlage von fachlich-disziplinärem Sachverstand und einschlägigen Urteilskompetenzen, unmittelbar in Forschungsprozessen zum Zwecke der Datenanalyse und -auswertung, der Mustererkennung usw., oder im Kontext wissenschaftlicher Praxis, etwa bei Recherchearbeiten und Textzusammenfassungen, bei der Erstellung von Forschungsplänen, bei der Erstellung von Programmcodes, bei Entwürfen und Gestaltungsaufgaben sowie für Aufgaben in der Wissenschafts-

kommunikation. Dabei dürften sich all diejenigen Aspekte auch langfristig nicht an KI delegieren lassen, die mit dem menschlichen in-der-Welt-sein verflochten sind: die Artikulation von Erkenntnisinteressen (warum wollen wir etwas wissen?), der kritische Diskurs über die Geltungsbedingungen wissenschaftlichen Wissens und die anthropologische (und dabei wohl auch anthropozentrische) Dimension von Erkenntnis.

3. KI und die Bildung der Person

Unter Bildung sei eine spezifische Entwicklung von Welt-, Sozial- und Selbstverhältnissen verstanden. Der Bildungsbegriff veranschlagt dabei die philosophischen Konzepte des Subjekts und der Person und prämiert die Fähigkeit zur verstehenden, selbstreflexiven und verantwortlich interessengeleiteten Auseinandersetzung mit der eigenen Verwobenheit in diejenigen Sachverhalte, die von individueller und kollektiver Bedeutsamkeit sind.

Zentrale Momente des Bildungsbegriffs sind verstehendes Aneignen, gemeinsam geteilte Bedeutsamkeit und Teilhabe. Dabei erfolgt verstehendes Aneignen nicht nach einem Zuschauermodell, sondern durch Erfahrungen im tätig-operativen Umgang und eine anschließende Reflexion und Interpretation dieser Erfahrungen. Diese Verkoppelung von Verstehen und Handeln beruht auf einem genetischen Prinzip: An Wissens- und Könnensordnungen individuell anzuschließen erfordert, diese nicht einfach als propositionale Aussagenkomplexe oder Verhaltensvorschriften aufzufassen, sondern als Resultate einer vorgängigen Praxis, die nach- und mitzuvollziehen ist – und zwar nicht lediglich rezeptiv, sondern eigenaktiv und reflektiert. Bildung erschöpft sich aber nicht im Verstehen, Bildung zielt auch auf Verfügungserweiterung und Teilhabe. Gemeinsam geteilte Bedeutsamkeit entsteht dann in dialogischen Prozessen zwischen Personen mit biographischen Erfahrungen, Weltverständnissen, Haltungen, Zielvorstellungen u.a.m.

Begreift man den Bildungsbegriff als eine Artikulationsform reflexiv-anthropologischer Selbstvergewisserung und (normativer) menschlicher Selbstauslegung, so fokussiert Bildung die individuelle und kollektive Verarbeitung der conditio humana, die zwar subjektbezogen erfolgt, aber auch als transpersonale Disposition zur intersubjektiven hermeneutischen, kritischen und reflexiven Verarbeitung von Weltzugängen und Weltaufschlüssen aufgefasst werden kann.

Wie verändert sich Bildung durch generative KI? Zunächst dürften sich durch KI weniger Bildungsprozesse als vielmehr die zugrunde liegenden Lernprozesse verändern. Die Nutzung von KI bedeutet technisch die Nutzung eines maschinellen Daten- und Informationsmanagementsystems; Wissen und Können dagegen erfordern eine Person, die zur persönlich-leiblichen Erfahrung des selber Machens (einschließlich des selber Denkens) fähig ist, welche sie in der Auseinandersetzung mit dem jeweiligen Gegenstand und seiner Widerständigkeit und Gestaltbarkeit gewinnt. Ob mit oder ohne KI – Daten und Informationen müssen in personales Wissen transformiert werden; und Tiefenwissen wiederum ist zentral, um neues Wissen einordnen zu können. Die Verwendung von KI im Zuge von Lernprozessen erfolgt idealerweise auf der Grundlage eines Wissens über KI und ihre Funktionsprinzipien. Die Nutzung von KI als technischem Artefakt erfordert allerdings nicht (nur) wie bei einem instrumentellen Werkzeuggebrauch die handwerkliche Beherrschung

des Werkzeugs und auch nicht (nur) wie bei der Bedienung einer Maschine ein Wissen für den Umgang mit dieser Maschine zum Zwecke ihrer geschickten Handhabung. Die Verwendung von teilautonomen Assistenzsystemen erfordert, ihren Output in den Kontext seiner technischen Herstellung einzuordnen, zu beurteilen, zu prüfen und kritisch zu reflektieren.

Fragt man darüber hinaus nach dem Verhältnis von *Bildung* und KI, so ist Gegenstand von Bildung (im oben dargelegten Verständnis), was sich in der Interaktion mit KI unter den Kriterien humaner Welt-, Sozial- und Selbstverhältnisse herausbildet bzw. herausbilden kann. KI könnte hier ein Auslöser für eine anthropologische Selbstreflexion und Selbstbesinnung sein, die im Bildungsbegriff ihren Resonanzraum findet: Der Bildungsidee wohnt eine (normative) Idee des genuin Menschlichen inne, sie artikuliert stets eine Selbstauslegung „des" Menschlichen als Gattungseigenschaft, das durch Bildung entfaltet wird. Insofern ist „humane Bildung" ein Pleonasmus, der allerdings die Abgrenzung zum Nicht-Humanen noch einmal explizit hervorhebt. In der Geschichte des Bildungsbegriffs waren die Abhängigkeit von Autoritäten, damit die Nicht-Selbstverfügung, und Prozesse der Entfremdung nicht-humane Momente, auf die die Bildungsidee mit dem Konzept der Bildung des Subjekts geantwortet hat. Aktuell könnte es die KI sein – gar in der materiellen Gestalt eines androiden Roboters, als nicht-menschlicher Ko-Akteur, als Homunkulus und technische Realisierung von „other minds" (wohl wissend, dass es sich lediglich um schwache KI handelt) –, auf die die Bildungsidee nunmehr mit einem Konzept der Bildung der Person reagieren könnte.

4. Die Idee(n) der Universität

Universitäten sind hybride Institutionen, in denen Forschung und Lehre programmatisch verkoppelt werden (hinzu kommt als dritte Funktion der Wissens- und Technologietransfer als Austausch mit außeruniversitären Akteuren). Über die vielfältigen Formen der Erzeugung und Weitergabe wissenschaftlichen Wissens und deren Zwecke ist damit aber noch nichts ausgesagt; es sind dann die verschiedenen Ideen der Universität, die diesem Komplex eine konzeptionelle Ausdrucksgestalt geben. Die Idee der Universität als Institution ist dabei von der empirischen Verfasstheit der Universität als Organisation zu unterscheiden. Insofern stellt jede Idee der Universität zunächst einen Idealtypus dar: Erstens sind auch Universitäten insbesondere „soziale Gebilde, in denen Menschen ihren individuellen Strategien folgen (…). Die intellektuellen Kriterien der Institution können sehr lebenspraktische Wünsche bemänteln. Anders als die institutionelle Leitidee es in immer gewandelter Form nahelegte, tendieren auch Universitäten zur Erstarrung und zum innovationsfeindlichen Dogmatismus." (Stölting, 2001, S. 34). Zweitens sind die Hochschulen verkoppelt mit außerhochschulischen wissenschaftlichen und außerwissenschaftlichen Tätigkeitsstrukturen und Arbeitsfeldern (die „Hochschule der Gesellschaft"). Als Bildungseinrichtungen haben sich die Hochschulen schon immer mit bildungspolitischen Steuerungsimpulsen und externen normativen Ansprüchen konfrontiert gesehen (vgl. exemplarisch Krücken & Musselin, 2009).

Vor diesem Hintergrund enthält jede Idee der Universität einen idealistischen Überschuss. Hält man dennoch an der Idee der „Idee der Universität“ (sic) fest, ist eine heuristische Zuspitzung denkbar: In dieser Auslegung ist die Universität nicht lediglich Lehranstalt zur Vermittlung von Wissen und akademischen Kompetenzen für praktische Anliegen, sondern ein Denk- und Reflexionsraum. Diese Universität ist als „transgressive Lebensform“ „ein Ort, an dem Begriffe und Einsichten beständig neu zu denken sind und neu gedacht werden können“ (Mecheril & Klingler, 2010, S. 84); die Universität ist ein Ort, „an dem nichts außer Frage steht“ (Derrida, 2001, S. 14). In diesem Sinne stellt eine solche unbedingte Universität radikale Fragen. Sie verlangt, Unsicherheit auszuhalten und ist sowohl Zu-Mutung als auch Zumutung.

Etwas nüchterner ließe sich als Idee der Universität und des akademischen Studiums ansetzen, dass die Studierenden durch ein wissenschaftliches Studium für die verantwortliche Bearbeitung komplexer Frage- und Problemstellungen substantiell *disponiert*, aber (noch) nicht für konkrete Berufs- und Arbeitsfelder *qualifiziert* werden. Es geht dann im Studium zunächst um die Befähigung für ein forschend-erkundendes, reflexives und methodisch basiertes begriffliches, konzeptuelles und materiales verstehend-erklärendes Modellieren, Erarbeiten oder Entwerfen – und dieses als Prozess der Entwicklung von Antwortvorschlägen zu komplexeren, ergebnisoffenen Fragestellungen, wobei nicht nur bestehendes Wissen einfließt, sondern auch neues Wissen im Prozess des Entwerfens von Antwortvorschlägen und Lösungsideen erarbeitet wird. Dabei spielt insbesondere auch der Umgang mit „verzwickten, bösartigen“ Problemen (Gänshirt, 2007, S. 18; vgl. auch Rittel, 1992) eine Rolle, die sich aufgrund ihrer Komplexität und Widersprüchlichkeit weder eindeutig bestimmen noch vollständig lösen lassen, die aber zugleich unaufgebbar sind, weil sie grundlegende Fragen berühren. Die Bearbeitung solcher komplexeren Frage- und Problemstellungen erfordert Verantwortungsbewusstsein, Sachkunde, Sachverstand, Urteilskompetenz und Kreativität, die im Studium (weiter)entwickelt werden.

Dabei darf der Einsatz von KI im Studium nicht dazu führen, durch die Delegation von anspruchsvollen Aufgaben an die KI intellektuelle Abkürzungen zu nehmen, indem durch die KI zwar Aufgaben bearbeitet werden (vorbehaltlich der Qualität der Ergebnisse dieser Bearbeitungen), sich dabei aber keine personale Sachkunde und kein Sachverstand entwickelt. Neben dem kontrollierten Einsatz von RAG-Chatbots[1] als KI-Tutoren dürften daher auch KI-freie Räume eine zentrale Rolle spielen.

5. Wissenschaft und Bildung – und KI

Inwiefern lassen sich (die) Relevanzdimensionen der lernenden, performativen und reflexiven Befassung mit Wissenschaft und ihren Verwendungsoptionen bzw. der Teilhabe an Wissenschaftspraxis mit dem Begriff der (akademischen) Bildung konzeptualisieren? Wie also hängen Wissenschaft und Bildung zusammen?

[1] Retrieval-Augmented Generation (RAG) bezeichnet technische Systeme, die auf Informationen aus spezifizierten Datenquellen zugreifen und nicht nur auf die Trainingsdaten des Large Language Modells.

Inwiefern werden diese Verhältnisse von Wissenschaft und Bildung durch generative KI beeinflusst? Gefragt wird damit nach Bildungsprozessen in der unmittelbaren Ausübung von Wissenschaft unter Einsatz von KI; nach der Einordnung von wissenschaftlichen Erkenntnissen, die unter Zuhilfenahme von KI gewonnen wurden; nach der Reflexion auf das epistemische Subjekt, das sich mit der Eigenlogik von digitalem Wissen und mit maschineller Epistemologie auseinandersetzt – im Kontrast und als Komplement zu menschlichem Wissen; nach den intellektuellen Leistungen des akademischen Subjekts, dem „Anderen der Vernunft“ und Mensch-Maschine-Verhältnissen im Kontext von KI.

6. … im 21. Jahrhundert

Die Außenseite einer Zeitdiagnose ist ihre historische Kontextualisierung. Als Figuren der historischen Einordnung lassen sich – ohne Anspruch auf Trennschärfe – drei Ebenen unterschiedlicher historischer Geschwindigkeiten veranschlagen: Ereignisse, Konjunkturen und Prozesse von langer Dauer (vgl. dazu Braudel, 1977; explizit mit dieser Theoriefigur arbeiten in erziehungswissenschaftlichem Kontext Forneck & Wrana, 2005).

Ereignisse sind in diesem Sinne etwa der Bologna-Prozess und die Studienstrukturreformen mit ihren weit reichenden Veränderungen. Zu Ereignissen zählen auch die Entwicklung und Implementierung neuer hochschuldidaktischer Gestaltungsformen, und Ereignisse sind einzelne technische Entwicklungen, etwa der generativen KI.

Konjunkturen als „Prozesse von langsamerer Geschwindigkeit“ (Forneck & Wrana, 2005, S. 7) zeigen sich beispielswiese in der Abfolge hochschulpolitischer Leitgedanken und Schlüsselthemen (etwa Fokussierung auf Kompetenzorientierung, Bildung für nachhaltige Entwicklung u.a.m.), in der wiederholten Rezeption oder in der zeitweiligen Dominanz bestimmter paradigmatischer Ansätze (z.B. empirisch-quantitative Forschung, reformpädagogische Argumentationen, Prominenz konstruktivistischer Lerntheorien). Konjunkturen verweisen aber auch auf Veränderungsprozesse in der Hochschullandschaft, etwa neue Steuerungsmodelle, Leistungsvergleiche, Verfahren der internen und externen Evaluation und Qualitätssicherung (Friebertshäuser, 2004, S. 248). Zu Konjunkturen gehören insofern Tendenzen der Metrisierung und Vermessung (Mau, 2017) und die Etablierung von Evaluationsregimen (vgl. auch Krüger, 2022; Reinbacher, 2019). Und Konjunkturen finden sich nicht zuletzt im Zusammenhang mit realgeschichtlichen Kontexten, etwa Autoritätsverluste wissenschaftlichen Wissens (Schäfer & Thompson, 2011, S. 23) oder Wandlungsprozesse der Wissenschaft von der Grundlagenforschung zu einer verstärkten Anwendungs- und Verwendungsorientierung (vgl. Carrier, 2006, S. 149 ff.).

Prozesse langer Dauer, aus denen heraus sich Konjunkturen erklären lassen (Forneck & Wrana, 2005, S. 7) verweisen schließlich auf realgeschichtliche und politische, gesellschaftliche und kulturelle Verhältnisse, auf Mentalitäten, politische Kultur und Dispositive, auf Veränderungen der Dominanz von Funktionssystemen und auf Transformationen gesellschaftlicher Praktiken; auf grundlegende Leitideen und Narrative und Veränderungen intellektueller Deutungsmuster. Bedeutsam für den Stellenwert von Universitäten und wissenschaftlicher Bildung könnten u.a. sein:

- die Amalgamierung von Globalisierung und Ökonomisierung;
- umfassende gesellschaftliche Veränderungsprozesse im Zuge von Automatisierung und Digitalisierung; damit verbunden die ideologischen Entwicklungslinien eines Trans- und Posthumanismus;
- Fragen der Nachhaltigkeit im Anthropozän;
- die postmoderne Freisetzung des Subjekts im Zuge sozio-kultureller Pluralisierung und seiner gleichzeitigen Gefährdung im Kampf um Anerkennung (Honneth, 1992) von Differenz, Diversität und Alterität;
- Herausforderungen an Demokratie und Zivilgesellschaft.

7. Und was ist das Fazit?

Wissenschaft ist eine spezielle epistemische und soziale Praxis im Kontext menschlicher Gesamtpraxis. Universitäten sind Organisationsformen für Orte zur Erzeugung und Weitergabe wissenschaftlichen Wissens (Weber, 2002, S. 9); sie übernehmen damit bestimmte Funktionen im Hinblick auf menschliche Gesamtpraxis und deren historisch-gesellschaftlich-kulturelle Spezifizierung und Ausdifferenzierung sowie im Hinblick auf individuelle (Aus-)Bildungsprozesse. Die Universität ist immer die Universität der Gesellschaft, die Idee der Universität entsteht in der Koevolution von Wissenschaftsgeschichte, Universitätsgeschichte, Technik- und Mediengeschichte und Praxisfeldern (z.B. Professionen) und den damit verbundenen Veränderungen gesellschaftlicher Wissensordnungen.

Vielleicht bleiben am Ende vor allem drei Fragenkomplexe:

- In welchem gesellschaftlichen Funktionszusammenhang steht wissenschaftliche Bildung? Und welche Rolle spielt KI in diesem Funktionszusammenhang?
- In welchem Ausmaß ist eine artikulierte Idee der Universität auch ein möglicher Realtypus, und inwieweit bleibt sie lediglich ein Idealtypus? Welche Hoffnungen, Sehnsüchte und Gegenentwürfe artikulieren sich in einer solchen Idee?
- Inwiefern ist die diskursive Verhandlung einer konkreten Idee der Universität lediglich ein Ereignis, inwiefern ist sie Ausdruck einer Konjunktur oder Ausdruck einer Kritik an einer Konjunktur, und inwiefern ist sie explizit oder implizit verwoben mit dem derzeitig andauernden Prozess langer Dauer? Und befinden wir uns aktuell an einem Wendepunkt zwischen Ereignissen, zwischen einer alten und einer neuen Konjunktur oder sogar zwischen zwei Prozessen langer Dauer?

Literatur

Braudel, F. (1977). Geschichte und Sozialwissenschaften. Die longue durée. In M. Bloch, F. Braudel, L. Febvre u.a., Schrift und Materie der Geschichte. Vorschläge zur systematischen Aneignung historischer Prozesse. Frankfurt / M.: Suhrkamp.

Carrier, M. (2006). Wissenschaftstheorie zur Einführung. Hamburg: Junius.

Derrida, J. (2001). Die unbedingte Universität. Frankfurt / M.: Suhrkamp.

Forneck, H. J. & Wrana, D. (2005). Ein parzelliertes Feld. Eine Einführung in die Erwachsenenbildung. Bielefeld: Bertelsmann.

Friebertshäuser, B. (2004). Hochschulforschung. In H.-H. Krüger & C. Grunert (Hg.), Wörterbuch Erziehungswissenschaft. Wiesbaden: VS Verlag für Sozialwissenschaften (S. 244 – 249).

Gänshirt, C. (2007). Werkzeuge für Ideen. Einführung ins architektonische Entwerfen. Basel u. a.: Birkhäuser.

Gramelsberger, G. & Sigwart, H.-J. (2025). Epistemologien zur Einführung. Hamburg: Junius.

Honneth, A. (1992). Kampf um Anerkennung. Frankfurt / M.: Suhrkamp.

Kornwachs, K. (2013). Philosophie der Technik. München: Beck.

Krücken, G. & Musselin, C. (2009). Europa und europäischer Bildungsraum. In S. Andresen u. a., Handwörterbuch Erziehungswissenschaft. Weinheim, Basel: Beltz (S. 263 275.).

Krüger, A. K. (2022). Soziologie des Wertens und Bewertens. Bielefeld: transcript.

Mau, S. (2017). Das metrische Wir. Über die Quantifizierung des Sozialen. Berlin: Suhrkamp.

Mecheril, P. & Klingler, B. (2010). Universität als transgressive Lebensform. In L. Darowska, T. Lüttenberg & C. Machold (Hg.), Hochschule als transkultureller Raum? Kultur, Bildung und Differenz in der Universität. Bielefeld: transcript (S. 83 – 116).

Nuxoll, F. (2024). Skill Skipping. Was Lehrpersonen tun können, damit KI nicht Lernprozesse verhindert. Lehren & Lernen 50 (6), S. 11-12

Reinbacher, P. (2019). Qualität und Qualitätsmanagement im Universitäts- und Hochschulbetrieb: Plädoyer für einen Paradigmenwechsel. Weinheim, Basel: Beltz Juventa.

Reinmann, G. (2023). Deskilling durch Künstliche Intelligenz? Potenzielle Kompetenzverluste als Herausforderung für die Hochschuldidaktik. Diskussionspapier. https://www.heise.de/news/Deskilling-Kompetenzverlust-durch-KI-wird-zu-wenig-diskutiert-9342124.html [

Rittel, H. W. (1992). Planen, Entwerfen, Design. Ausgewählte Schriften zu Theorie und Methodik. Herausgegeben von W. D. Reuter. Stuttgart u.a.: Kohlhammer.

Schäfer, A. & Thompson, C. (Hg.). (2011). Wissen. Paderborn: Schöningh.

Stölting, E. (2001). Permanenz und Veränderung von Strukturkrisen: Institutionelle Darstellungsprobleme. In E. Stölting & U. Schimank (Hg.), Die Krise der Universitäten. Wiesbaden: Westdeutscher Verlag (= Leviathan Sonderheft 20).

Weber, W. E. J. (2002). Geschichte der europäischen Universität. Stuttgart: Kohlhammer.

Wulf, C. (2020). Bildung als Wissen vom Menschen im Anthropozän. Weinheim, Basel: Beltz Juventa.

Rittel, H. W. (1992). Planen, Entwerfen, Design. Ausgewählte Schriften zu Theorie und Methodik. Herausgegeben von W. D. Reuter. Stuttgart u.a.: Kohlhammer.

Schäfer, A. & Thompson, C. (Hg.). (2011). Wissen. Paderborn: Schöningh.

Stölting, E. (2001). Permanenz und Veränderung von Strukturkrisen: Institutionelle Darstellungsprobleme. In E. Stölting & U. Schimank (Hg.), Die Krise der Universitäten. Wiesbaden: Westdeutscher Verlag (= Leviathan Sonderheft 20).

Weber, W. E. J. (2002). Geschichte der europäischen Universität. Stuttgart: Kohlhammer.

Wulf, C. (2020). Bildung als Wissen vom Menschen im Anthropozän. Weinheim, Basel: Beltz Juventa

Von der Schockstarre zur Symbiose?! Acht Gespräche über generative KI in der Hochschule

Andreas Hebbel-Seeger

Im Rahmen des PodCasts „LectureCast“ spreche ich seit vielen Jahren mit Hochschullehrenden über wechselnde Themen von hochschuldidaktischer Relevanz (Hebbel-Seeger, i.D.). Generative KI im Kontext des Lehrens, Lernens und Forschens an Hochschulen war erstmals Anfang 2023 Gegenstand eines Gesprächs, das ich mit Doris Weßels geführt habe; nur wenige Wochen nach der Veröffentlichung von ChatGPT durch das US-amerikanische Unternehmen OpenAI, der ersten frei zugänglichen Chat-Oberfläche zu ihrem Sprachmodell GPT-3.5. Seitdem ist das „KI-Thema“ nicht nur im öffentlichen wie im akademischen Diskurs zu einem allgegenwärtigen Inhalt geworden, sondern es durchzieht auch zunehmend die Gespräche in „meinem“ Podcast; selbst dann, wenn die thematischen Setzungen jeweils „eigentlich“ andere sind. Gleichwohl habe ich generative KI in der Folge des eben genannten Gesprächs mit Doris Weßels bis heute in weiteren acht Podcast-Gesprächen explizit zum zentralen Gesprächsinhalt gemacht, wobei ich eines davon in Episode 101 im „Dialog“ mit den KI-Instanzen ChatGPT, Claude und Gemini geführt habe.[1]

Die im Folgenden betrachteten Episoden dokumentieren exemplarisch die Entwicklung des hochschuldidaktischen Diskurses über generative KI. Sie reichen von ersten Reaktionen auf ein als disruptiv wahrgenommenes Werkzeug bis hin zu aktuellen Diskussionen über KI-Agenten, institutionelle Leitlinien und neue Kompetenzanforderungen. Trotz unterschiedlicher Perspektiven zeigen sich dabei Gemeinsamkeiten hinsichtlich der Herausforderungen und Chancen für Lehre, Lernen und Prüfungskultur, was wiederum auch die Schnittstelle zu der hier vorliegenden Essaysammlung bildet.

Disruption als Ausgangsdiagnose

Doris Weßels (damals FH Kiel) beschreibt in der LectureCast Episode 57[2] den durch ChatGPT ausgelösten Wandel als tiefgreifende Veränderung akademischer Schreibprozesse: Studierende produzieren Texte nicht länger allein, sondern zunehmend gemeinsam mit digitalen Systemen. Als Konsequenz dieser Entwicklung schlägt Weßels aber nicht eine (ohnehin auch damals schon nicht technisch zuverlässig zu leistende) Kontrolle der studentischen Arbeitsergebnisse vor, sondern eine (Selbst-)Überprüfung der Aufgabenstellungen durch die Lehrenden mittels generativer KI: Liefert die KI dabei überzeugende Ergebnisse, müsse die Aufgabenstellung überdacht werden. Diese Perspektive prägt viele der späteren Gespräche. Gleichzeitig verweist Weßels auf ein Szenario, das die Debatten bis heute begleitet: Studierende könnten Texte von KI-Systemen erstellen lassen, während Lehrende diese wiederum mithilfe von KI bewerten. Die eigentliche Sorge richtet sich dabei weniger

[1] https://open.spotify.com/episode/5kFIqI0ICWKdBcZNCsh8W3?si=qDRzhb-CT6SjUiIOCgpMaw
[2] https://open.spotify.com/episode/11OQbI8WQ5hc1dgAsKy7EI?si=TJePwtyYTSW-VZ1VGGwIzQ

auf die Technologie als auf die Gefahr, dass zentrale Lern- und Bewertungsprozesse zunehmend ausgelagert werden.

Wenige Monate später berichtet Dirk Kleine (Macromedia University) in Episode 66[3] des LectureCast von ersten praktischen Erfahrungen in der Lehre. Seine Beispiele zeigen, wie rasch aus abstrakter Antizipation konkrete Praxis wurde: Tools wie ChatGPT lieferten im Prompting-Dialog vollständige Veranstaltungskonzepte samt Literaturhinweisen und in einem interdisziplinären Projekt entwickelten Studierende mit Hilfe generativer KI Geschäftsideen und Pitch-Videos. Ausgehend von eigenen, auch außeruniversitären Erfahrungen argumentiert er, dass generative KI Kreativität nicht einschränke, sondern gerade Menschen ohne einschlägige Vorbildung eine kreative Beteiligung oft erst ermögliche. Routinetätigkeiten würden abgenommen, sodass mehr Kapazität für gestalterische Entscheidungen bleibe. Für die Lehre folgert Kleine, dass es weniger eines eigenständigen KI-Curriculums als vielmehr eines fächerübergreifenden, verantwortungsvollen Umgangs mit der Technologie bedürfe – eine Position, die sich später in institutionellen Leitlinien (z.B. an der Universität Heidelberg oder der Universität Hamburg), Whitepapers (z.B. Mütterlein et al., 2024) und nicht zuletzt auch in einer Empfehlung des Wissenschaftsrats (2026) wiederfindet.

Von der Ergebnis- zur Prozessbewertung im Prüfungswesen

Wie sich diese Entwicklung auf Prüfungsformate auswirkt, wird im Gespräch mit Joschka Mütterlein (Macromedia University) in Episode 70[4] des LectureCast deutlich. Er vergleicht die Bedeutung generativer KI für wissenschaftliche Arbeiten mit jener des Taschenrechners für mathematische Prüfungen: Nicht das Prüfungsformat als solches wird obsolet, wohl aber die bisherige Form. Aus seiner Beschreibung der institutionellen Reaktion der Hochschule lässt sich ablesen, wie folgenreich diese Verschiebung sein kann: Die Macromedia University entschied sich vergleichsweise früh, den Einsatz generativer KI durch Studierende nicht nur zu erlauben, sondern aktiv zu fördern und zu fordern, was nach Mütterlein zur Ausbildung von vier unterschiedlichen Nutzungsprofilen unter den Studierenden geführt hat: (a) jene, die KI transparent und reflektiert nutzen, (b) jene, die sie verdeckt einsetzen, (c) jene, die sie aus Unsicherheit meiden, und schließlich (d) jene, an denen die Entwicklung bis zu diesem Zeitpunkt (gut 1,5 Jahre nach dem Launch von ChatGPT) vorbeigegangen sei. In einer ersten Konsequenz wurde die Bewertungslogik von Projektarbeiten an der Macromedia University angepasst: Die Gewichtung von Form und Struktur wurde gesenkt, während mit der Prozessreflexion ein neuer Bewertungsaspekt eingeführt wurde. Damit wurde hier zumindest formal vollzogen, was in vielen folgenden Gesprächen als Grundmuster wiederkehrt: die Verschiebung des didaktischen Blicks vom fertigen Produkt auf den Weg dorthin.

[3] https://open.spotify.com/episode/1gLUqL1A5K0RewBry9NoKu?si=fMxjffVvRfm_KMN0Q7qxsw
[4] https://open.spotify.com/episode/2lHO1PUZxZfydJ9DSfFGsN?si=5DI_yEpaR1GyJ8e6Iuwbfg

Disruption als Dauerzustand

Als ein weiteres Motiv durchzieht die Mehrheit der Gespräche die hohe Dynamik der technologischen Entwicklung. Neue Werkzeuge, Funktionen und Anwendungsfelder entstehen in einer Geschwindigkeit, die klassische Fortbildungsformate kaum abbilden können. Dirk Kleine beschreibt dies als permanente Aktualisierungsaufgabe. Ähnliche Beobachtungen finden sich bei Elisabeth Mayweg (Humboldt-Universität Berlin) in der LectureCast Episode 120[5] und Marcus Birkenkrahe (Augustana University, South Dakota/US) in LectureCast Episode 126[6]. Beide betonen die Bedeutung informeller Lernnetzwerke, kollegialen Austausches und experimenteller Lehrformate: Lehrende befinden sich häufig selbst in Lernprozessen und erkunden neue Werkzeuge gemeinsam mit ihren Studierenden. Institutionelle Unterstützung kann sich unter diesen Bedingungen nicht auf einzelne Leitfäden oder Schulungen beschränken. Vielmehr benötigt die Hochschullehre, so meine Interpretation, kontinuierliche Austausch- und Lernräume, welche eine Reaktion und Orientierung an der Dynamik der technologischen Entwicklung möglich machen.

Wissenschaftliches Schreiben zwischen Unterstützung und Delegation

Besonders deutlich zeigen sich die Auswirkungen generativer KI beim wissenschaftlichen Schreiben. Barbara Geier (FH Burgenland, AUT) weist in LectureCast Episode 88[7] darauf hin, dass KI-generierte Texte begrifflich nicht als Plagiate verstanden werden können, da kein ursprünglicher Referenztext kopiert wird. Treffender sei die Analogie zum Ghostwriting. Als Reaktion gewinnen Prüfungsformate an Bedeutung, die schriftliche Arbeiten durch mündliche Elemente ergänzen. Gleichzeitig werden Portfolios, projektorientierte Arbeiten und andere praxisnahe Formate zunehmend als Alternativen zur klassischen Hausarbeit diskutiert. Geier betont explizit die Potenziale generativer KI: Menschen ohne muttersprachliche Englischkenntnisse seien im internationalen Wissenschaftsbetrieb nicht mehr in gleichem Maße benachteiligt und komplexe Inhalte ließen sich für ein breite(re)s Publikum verständlich machen. Entscheidend sei, so Geyer, dass Studierende die Vorschläge der Systeme kritisch beurteilen können. Problematisch werde der Einsatz dort, wo die fachliche Kontrolle über die Ergebnisse verloren gehe: Wer KI nicht als Co-Pilot nutze, sondern selbst zum Co-Piloten der KI werde, also unreflektiert Textvorschläge übernehme, ohne die fachliche Substanz beurteilen zu können, stellt wissenschaftliche Integrität in Frage.

Ausgehend von kognitionswissenschaftlichen Überlegungen argumentiert Dagmar Schulze-Heuling (Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg) in Episode 117[8] des LectureCast, dass Schreiben und Denken interdependent sind: Gedanken würden nicht fertig aufgeschrieben, sondern erst im Schreibprozess selbst geformt; ein Prozess des „Knowledge Transformings“ statt eines bloßen „Knowledge Tellings“. Wer das Schreiben an eine KI delegiere, verlagere damit auch einen Teil des Denkens, was Schulze-

[5] https://open.spotify.com/episode/4DOUjbwHVRohCXumTUi5ew?si=ChH7XwwSSx6adMU4Zx08Pg
[6] https://open.spotify.com/episode/5lvZh96Z4yyvGgQeWsBaet?si=mqKWKlPVQLCLArLGG-aUew
[7] https://open.spotify.com/episode/5t7vlQ2SBqT62CfgHuQImu?si=bUFKMUsDQtaf_5Od1vY8fQ
[8] https://open.spotify.com/episode/508NBg2Pq4PqxVGiUTZAYy?si=2jg5ODozStW7HzTK16hmqA

Heuling unter Verweis auf empirische Studien als “kognitives Offloading” bezeichnet (ähnlich äußern sich auch Herzberg & Reinmann (2025, S. 3) die ein „Conversational Offloading“ als einen von vier Interaktionsmodi im Dialog mit einem Sprachmodell skizzieren). Einen möglichen Lösungsansatz sieht Schulze-Heuling in einer Orientierung an der Bloomschen Lernzieltaxonomie: Auf den unteren Kompetenzstufen Wissen und Erinnern hält sie den Einsatz generativer KI für unproblematisch, auf den Stufen der Anwendung und Analyse legt sie Zurückhaltung nahe, während sie (anders als bspw. vom Wissenschaftsrat (2026) vertreten, der diesbezüglich, orientiert am Leitbild einer „intellektuellen Souveränität“, keine Einschränkungen macht) die Nutzung generativer KI auf den Stufen des Bewertens und Kreierens als hochproblematisch einschätzt. Als zentrale Gefahr benennt Schulze-Heuling, den Co-Autor einer gemeinsamen Publikation zitierend, in diesem Sinne nicht die Technologie selbst, sondern eine aus seiner unreflektierten Nutzung resultierende „algorithmischen Reproduktion parteiischer Gewissheit“ (Gerlich, 2025, S. 937) und die aus ihr erwachsende Bequemlichkeit, Inhalte nicht mehr kritisch zu hinterfragen.

KI-Kompetenz als Bildungsauftrag

Einen umfassenden Kompetenzrahmen beschreibt Elisabeth Mayweg in Episode 120[9]. Sie verortet AI Literacy in der Tradition der Medienbildung, betont jedoch die Besonderheit dialogischer KI-Systeme. Anders als klassische Informationswerkzeuge treten diese Systeme aktiv in Kommunikations- und Denkprozesse ein. AI Literacy umfasst für Mayweg daher drei Dimensionen: zum einen ein grundlegendes Verständnis der Technologie, zweitens die Fähigkeit zur kritischen Bewertung von Ergebnissen sowie drittens eine strategische und zielgerichtete Nutzung. Zusammengeführt werden diese Bereiche durch die Fähigkeit zu einem reflektierten Vertrauen, das weder von unkritischer Begeisterung noch von pauschaler Ablehnung geprägt ist. Ihre eigenen Beobachtungen in der Praxis zeigen indes ein heterogenes Bild: Viele Studierende, so Mayweg, nutzen KI vor allem zur Effizienzsteigerung. Andere setzen sie gezielt ein, um Perspektiven zu erweitern, Gegenargumente zu entwickeln oder Prüfungssituationen zu simulieren. Daraus ergibt sich eine neue Rolle für Lehrende, von Wissensvermittlern hin zu Lernbegleitern in einem gemeinsamen Erkundungsprozess. Gleichzeitig warnt Mayweg vor einer Verschärfung bestehender Bildungsungleichheiten. Wer bereits über Ressourcen, Vorwissen und Selbststeuerung verfügt, kann von KI-Systemen häufig stärker profitieren als Personen mit ungünstigeren Ausgangsbedingungen. KI schließt Bildungsunterschiede nicht automatisch, führt Mayweg aus, sondern kann sie ganz im Gegenteil unter Umständen sogar noch verstärken.

Vom Werkzeug zum Akteur

Die beiden jüngsten Gespräche verschieben den Fokus von der Textproduktion hin zu KI-Agenten und institutionellen Rahmenbedingungen. Marcus Birkenkrahe (Episode 126[10]) berichtet von gestuften KI-Richtlinien an seiner amerikanischen Hochschule, die von Level 0, keinem KI-Einsatz, bis Level 5, in dem KI-Einsatz erwartet wird, reichen, und

[9] a.a.O.
[10] a.a.O.

Lehrenden eine differenzierte, aufgabenspezifische Handhabung erlauben, anstelle pauschaler Freigaben oder Verbote. Birkenkrahe, der unter Verweis auf den Physiker und Nobelpreisträger Roger Penrose lieber von „künstlicher Klugheit“ als von „Intelligenz“ spricht, argumentiert mit prinzipiellen Grenzen der Technologie: Diese liegen in einem fehlenden Bewusstsein, dem unüberbrückbaren Sprung in die physische Welt, der Black-Box-Natur der Modelle, Halluzinationen und dem als „Context Drift“ bezeichneten Phänomen, dem schleichenden Abschweifen vom ursprünglichen Thema oder Ziel in der Konversation mit einer KI. Didaktisch übersetzt Birkenkrahe diese Skepsis in einen Code-Along-Modus, bei dem Studierende Programmieren nicht durch Zuhören, sondern durch gemeinsames, muskelgedächtnisbildendes Tun im Hörsaal erlernen, ergänzt um Scrum[11]-basierte Semesterprojekte, bei denen alle vier Wochen ein Prototyp präsentiert wird. Seine zentrale These lautet, dass gerade analoges, körperlich verankertes Wissen (im weitesten Sinne analog zu der von Frank Vohle in diesem Band skizzierten Handlungsebene „Erfahrung“), etwa über Hardware, an Bedeutung gewinnen wird, während komplexe, Ende-zu-Ende verantwortete menschliche Prozesse auf absehbare Zeit nicht ersetzbar seien.

Dominic von Proeck (Leaders of AI[12]) denkt in der LectureCast Episode 127[13] diese Perspektive weiter und beschreibt eine Zukunft, in der Mensch und KI zunehmend als kooperatives System agieren. Seine zentrale These lautet, dass die Kombination menschlicher und maschineller Fähigkeiten vielen Einzelakteuren überlegen sei. Voraussetzung bleibt jedoch eine fundierte fachliche Expertise: Nur wer über ausreichendes Domänenwissen verfügt, kann die Qualität von KI-Ergebnissen angemessen beurteilen. Für die Zukunft prognostiziert er daher, dass insbesondere soziale, kommunikative und führungsbezogene Kompetenzen an Bedeutung gewinnen werden, während gleichzeitig fachliches Wissen unverzichtbar bleibt.

Werte, Verantwortung und Verzerrungen

Die frühen Debatten dominierten vor allem Fragen nach der Leistungsbewertung. In den jüngeren Gesprächen rücken hingegen vor allem normative Aspekte stärker in den Blickpunkt. Schulze-Heuling verweist auf Untersuchungen zu politischen und kulturellen Schlagseiten unterschiedlicher Sprachmodelle sowie auf stereotype Darstellungen in Bildgeneratoren, die Diversität erst nach expliziter Aufforderung abbilden. Mayweg schließt daran mit ihrer Forderung nach Bias-Sensibilität als Bestandteil von AI Literacy an. Von Proeck wiederum berichtet aus unternehmerischer Perspektive, wie organisatorische Leitbilder und Prioritäten explizit in KI-gestützte Prozesse integriert werden können und sollten. D.h. Fairness und Nicht-Diskriminierung werden als Werte genannt, wenn es darum geht, mit (politischen und kulturellen) Verzerrungen in Sprachmodellen umzugehen, Repräsentation und Vielfalt als Orientierungsrahmen in Bezug auf Stereotype in Bild-

[11] Karsten Morisse (Hochschule Osnabrück) skizziert in LectureCast Episode 59 agiles Arbeiten in der Hochschullehre mit Verweis auf die aus der Softwareentwicklung stammende Scrum-Methode: https://open.spotify.com/episode/65TVNI7ZZNVqL7z91wdhMq?si=jTvnhGSxRjiGKXKaYbg8Vw

[12] Ein kommerzieller Weiterbildungsanbieter, der Führungskräfte und Organisationen bei der Integration künstlicher Intelligenz in Arbeits- und Lernprozesse unterstützt: https://www.leadersofai.com/

[13] https://open.spotify.com/episode/1tXNjfbZkVe4pLYod9AIpW?si=LulDZyo4SVS6IcxbTNZJEg

generatoren sowie organisationale Verantwortung, wenn Unternehmen die Integration von Leitbildern in KI-gestützte Prozesse angehen. Gemeinsam ist diesen Perspektiven die Überzeugung, dass ethische Fragen und eine damit zusammenhängende Werteorientierung nicht nachträglich behandelt werden dürfen, sondern bereits bei Entwicklung, Auswahl und Nutzung von KI-Systemen berücksichtigt werden müssen. Für die Hochschullehre bedeutet dies, dass KI-Kompetenz über technische Anwendungskompetenzen hinausgeht. Notwendig sind ebenso Kenntnisse über Quellenkritik, Transparenz, Verzerrungen, Verantwortung und die gesellschaftlichen Folgen algorithmischer Systeme. Diese Schlussfolgerung ist unmittelbar anschlussfähig an verschiedene Beiträge des hier vorliegenden Essay-Bandes, wobei sich allerdings nicht nur inhaltliche Parallelen, sondern auch Widersprüche zeigen: So betonen Steinhardt et al. (in diesem Band) ebenso wie Schulze-Heuling, dass KI-Systeme nicht neutral seien, sondern bestehende strukturelle Ungleichheiten verstärken oder abschwächen können, und fordern analog zu Mayweg eine kritische KI-Literacy, die über eine „bloße“ Anwendungskompetenz hinausgeht. Statt der von Schulze-Heuling beklagten Verzerrung der Ergebnisse von KI-Bildgeneratoren, nimmt Judith Simon im vorliegenden Band die Rezeptionsseite in den Fokus, in dem sie auf Personalisierung und Filterblasen in der Nutzung hinweist, welche gleichermaßen die Werte Freiheit und Gleichheit tangieren. Während von Proeck für die Integration organisationaler Leitbilder in KI-gestützte Prozesse plädiert, hält Ines Langemeyer (in diesem Band) eine derartige Beteiligung generativer KI für unverantwortlich, weil, wie sie ausführt, nur der Mensch und eben nicht Maschinen in eine reflexive Auseinandersetzung mit Werten wie Gleichheit oder Freiheit treten kann. Aus der Erwartung, Werte (auch) technisch verankern zu können, und dem Postulat einer ausschließlich dem Menschen zueigenen Verantwortung, leitet sich mithin ein Spannungsfeld ab, das basal für die zukünftige technologische Entwicklung einerseits aber eben auch für den praktischen Umgang damit andererseits sein dürfte.

Zusammenfassung und Ausblick

Trotz unterschiedlicher Zugänge, Perspektiven und Akzente lassen sich in den hier behandelten LectureCast-Gesprächen gemeinsame Muster erkennen: Generative KI ist gekommen, um zu bleiben. Sie ist und wird Bestandteil von Hochschullehre und Wissenschaft sein. Die zentrale Frage lautet daher nicht, ob KI genutzt werden soll, sondern wie ihre Integration gestaltet werden kann. In diesem Zuge verschiebt sich der didaktische Fokus vom Ergebnis auf den Lernprozess, der wiederum, so die Essenz der hier zusammengetragenen Gespräche, weniger durch den Verlust von Kompetenzen gekennzeichnet ist als vielmehr durch eine Veränderung bzw. Verschiebung: An Bedeutung gewinnen Urteilsfähigkeit, Reflexionsvermögen, Domänenwissen und soziale Kompetenzen. Die Gefahr von Bildungsungleichheiten wurde wiederholt aber gleichwohl ambivalent thematisiert. Und schließlich zeichnet sich in den Gesprächen ab, dass nicht die Technologie selbst als zentraler Risikofaktor gesehen wird (nicht zuletzt im Unterschied zu aktuellen Äußerungen der Entwickler einschlägiger Plattformen, wie bspw. dem Anthropic-Mitgründer Amodei, 2026a, 2026b), sondern ein unkritischer und bequemer Umgang mit ihr.

Der Bogen von den ersten Reaktionen auf ChatGPT bis zu aktuellen Diskussionen über KI-Agenten lässt sich als Reifungsprozess eines hochschuldidaktischen Diskurses lesen. Aus anfänglicher Verunsicherung ist vielerorts eine differenzierte und reflektierte Praxis entstanden – gleichwohl, ohne dass grundlegenden Fragen beantwortet sind. Vielmehr entstehen mit jeder technologischen Entwicklung neue Herausforderungen für Lehre, Lernen und Prüfungskultur. Was sich bis heute allerdings nicht geändert hat, ist die Unsicherheit über den weiteren Verlauf. Wie es von Proeck formuliert: „Wir sind gerade alle gleichermaßen Schüler im Umgang mit KI."

Literatur

Amodei, D. (2026a). The Adolescene of Technology. Confronting and Overcoming the Risks of Powerful AI. https://darioamodei.com/essay/the-adolescence-of-technology [14.09.2026]

Amodei, D. (2026b). We Must Pace the Frontier. https://darioamodei.com/post/we-must-pace-the-frontier [14.09.2026)

Beratungskreis Digitalisierung in der der Lehre der UHH (2025). Orientierungsrahmen der UHH im Umgang mit generativen KI-Systemen in Studium und Lehre. https://www.uni-hamburg.de/lehre-navi/lehrende/orientierungsrahmen-gki/orientierungsrahmen-gki.pdf

Gerlich, M. (2025). Kognitives Offloading und die Erosion des kritischen Denkens im Zeitalter der generativen KI. In D. Schulze Heuling, A. P. Jakobi, G. S. Schaal, & M. Gerlich (Eds.), *Generative KI und (politik)wissenschaftliches Schreiben. Politische Vierteljahresschrift*, 66, 913–940. DOI: 10.1007/s11615-025-00631-9.

Hebbel-Seeger, A. (i.D.). Podcasts als Medium hochschuldidaktischer Weiterbildung. Nutzung, Akzeptanz und Gestaltungspräferenzen am Beispiel des „LectureCast" In n.n. (Hrsg.), *Hochschullehre im Wandel. Lebenslanges Lernen und akademische Praxis im Zeitalter von KI: Organisationale, soziale und didaktische Transformation*. Münster: Waxmann.

Herzberg, D. & Reinmann, G. (2025). Dialogmaschinen im Kommuniativen Beziehungsgeflcht der Hochschulbildung: Idealtypische Nutzungsmodi. *Impact Free*, 62. https://gabi-reinmann.de/wp-content/uploads/2025/03/Impact_Free_62.pdf

Mütterlein, J., Ranner, T. & Müller, M. (2024). Artificial Intelligence at Universities. Impact on Grades, Student Experience, and Teaching. Whitepaper, Macromedia University of Applied Sciences. München. https://macromedia-publications.bsz-bw.de/frontdoor/deliver/index/docId/487/file/AI_and_Universities.pdf

Universität Heidelberg (o.J.). Leitlinie zum Umgang mit KI in Studium und Lehre an der Universität Heidelberg. https://backend.heiskills.uni-heidelberg.de/de/dokumente/ki-leitlinie-de/download

Wissenschaftsrat (2026). Intellektuelle Souveränität: Empfehlungen für die Hochschulbildung in Zeiten von generativer KI. Köln. DOI: 10.57674/1evx-t906.

Teil I

Was KI mit Wirklichkeit und Erkenntnis macht

Die maschinelle Konstruktion der Wirklichkeit

Antje Michel

„Oberstes Ziel von Bildung ist es, den Menschen zu einer mündigen und freien, das heißt einer zur Verantwortung fähigen Person heranzubilden“ (Ethikrat 2023, S. 221), so wird es in der Stellungnahme „Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz“ des Deutschen Ethikrats formuliert. Dieses Ziel der „Persönlichkeitsbildung“ (Ethikrat 2023, S. 221) erfordere den Erwerb von Orientierungswissen aus der kritischen Auseinandersetzung mit empirischen Alltagserfahrungen und Forschungsergebnissen sowie aus der Positionierung zu normativen Inhalten. Bildung wird als Voraussetzung dafür betrachtet, dass der Mensch Verantwortung gegenüber seiner Umwelt und gegenüber sich selbst entwickelt. Bildung wird nicht allein durch die kognitive Informationsvermittlung ermöglicht, vielmehr bedarf es der sozialen Interaktion aller, die am Lehr-Lernprozess beteiligt sind. Noch vor dem Launch von ChatGPT im November 2022 empfahl der deutsche Ethikrat, Konzepte und Werkzeuge der digitalen Bildung dahingehend zu prüfen, ob sie diesen sozialen Interaktionsprozess fördern oder behindern. Der Wissenschaftsrat fächert das, was im allgemeinen Bildungsbegriff des Deutschen Ethikrates als Orientierungswissen zusammengefasst wurde, für die hochschulische Bildung in die Dimensionen (Fach-)Wissenschaft und Arbeitsmarktvorbereitung auf. Die hochschulische Ausprägung der Persönlichkeitsbildung wird als Entwicklung eines (fach)wissenschaftlichen Ethos sowie als Entwicklung eines kritisch-reflektierten gesellschaftsgestaltenden Rollenverständnisses spezifiziert (Wissenschaftsrat 2015, S. 9). Implizit wird auch hier die Relevanz der sozialen Interaktion für die Persönlichkeitsentwicklung betont, indem diese als (akademische) Sozialisation beschrieben wird (Wissenschaftsrat 2015, S. 9).

Sozialkonstruktivistische Theorien der Soziologie beschreiben die soziale Interaktion als eine aufeinander bezugnehmende Konstellation von (mindestens) zwei Beteiligten (Ego und Alter) und heben hervor, wie bedeutsam es ist, die Haltung des anderen antizipieren zu können. Das bedeutet: Soziale Interaktion zwischen zwei Beteiligten erfordert, dass Ego sowohl in der Lage ist, sich in die Position von Alter zu versetzen, und die Antizipation von Alters Reaktion in die Ausführung seiner Geste oder Äußerung einbezieht als auch davon ausgeht, dass Alter dasselbe tut und Ego somit vorhersehen kann, wie eine erwartbare Reaktion von Alter aussieht. Sofern es sich um symbolische Interaktionen handelt, antizipieren beide, dass ihre Gesten oder Äußerungen von Dritten ebenso antizipiert werden könnten, also generalisierbar sind. Dieses Prinzip der dreifachen Einstellungsübernahme (Lindemann 2025) dient sowohl als Ausgangspunkt für die Entwicklung von Theorien zur Identitätsentwicklung und Sozialisation im Symbolischen Interaktionismus von Mead und Blumer als auch – in einer dialektischen Beziehung zwischen Individuum und Gesellschaft – zur Aushandlung dessen, was wir als gesellschaftliche Wirklichkeit sozial konstruieren (Berger & Luckmann 1980). Hubert Knoblauch generalisiert in seiner Theorie des kommunikativen Konstruktivismus dieses Interaktions- bzw. Handlungsmodell als „wechselseitiges Wirkungshandeln“ (Knoblauch 2012, S. 6), indem er Habermas Unterscheidung von zweckrationalem und kommunikativem Handeln zusammenführt, wie auch

die Unterscheidung von körperlichem, nicht intentionalem, Verhalten und zielgerichteter, den Körper einbindender, Interaktion in Frage stellt. Ein so verstandenes kommunikatives Handeln bezieht sich stets nicht nur auf das handelnde Subjekt, sondern auch auf Andere und auf die mit dem Handlungsprozess verknüpften Objektivierungen, die – wiewohl kommunikativ konstruiert – den sozialen Kontext bzw. die gemeinsame Umwelt darstellen.

Die ‚Lockmimikry' generativer KI-Chatbots

Wenn kommunikatives Handeln wesentliches Prinzip von sozialer Interaktion ist und (Persönlichkeits-)Bildung in sozialer Interaktion erworben wird, stellt sich die Frage, was es bedeutet, wenn man generativer KI den Status eines Akteurs im sozialen Interaktionsprozess der Hochschulbildung zuweist. Meines Erachtens liegt das wesentlichste Problem nicht darin, dass generative KI eine Maschine ist, deren Interaktionsschemata anders sind als menschliche, sondern darin, dass generative KI diese Andersartigkeit maskiert und wir in der Interaktion einer fortwährenden Täuschung ausgesetzt werden, indem generative KI in der Interaktion ein menschliches Kommunikationsverhalten plausibel simuliert. Zum einen weisen bereits die ELIZA-Erfahrungen von Joseph Weizenbaum darauf hin, dass Menschen dazu neigen, einem Chatbot menschliche Eigenschaften wie Empathie oder Bewusstsein zuzugestehen, wenn dieser aus ihrer Sicht menschlich plausibel kommuniziert (vgl. Simon et al. 2024, S. 14). Zum anderen liegt meines Erachtens aber das Problem nicht nur bei den menschlichen Rezipient*innen, denn die aktuellen KI-Chatbots unterliegen einer Marktlogik und ihre Entwickler*innen haben ein Interesse daran, die menschlichen Nutzer*innen möglichst lange und möglichst zufrieden in einem Chat zu halten. Somit werden aktiv Strategien verfolgt, um menschliches Kommunikationsverhalten zu simulieren. Die Entwickler*innen von KI-Chatbots fördern bei ihrem Produkt eine Art kommunikative ‚Lockmimikry', indem sie die eigentliche Funktionslogik der Maschine mit Attributen menschlicher Kommunikation (wie z.B. Anreden, verständnissinnigen Floskeln, Emojis) ‚überschminken'.

Diese Simulation menschlicher Kommunikation kann die Funktionen von Kommunikation als menschlichem Instrument der Konstruktion von Identität und Welt durch kommunikatives Handeln nicht erfüllen. Menschliches kommunikatives Handeln bezieht stets den Körper (treffender noch: den Leib) mit ein, der die Handlung vollzieht und ihren Sinn sichtbar macht, und es vollzieht sich im Verlauf von Zeit (Knoblauch 2012).

Kommunikative Konstruktion von Wirklichkeit in der Hochschullehre

Was bedeutet es für hochschulische (Persönlichkeits-)Bildung, wenn Anteile der sozialen Interaktion im Lehr-Lerngeschehen an maskierte Maschinen ausgelagert werden und deren mimetische Simulation menschlicher Kommunikation an die Stelle des Originals tritt? Dies möchte ich in zwei Thesen skizzieren.

Meine erste These ist, dass dies dazu führen könnte, dass mit der Zeit unsere bisherigen Praktiken der (Persönlichkeits-)Bildung um wesentliche Merkmale reduziert werden, die Maschinen nicht leisten können: nämlich die Verknüpfung kognitiver, emotionaler und

leiblicher[1] Erfahrungen zur Entwicklung neuen Wissens oder Veränderungen von Perspektiven über die Zeit und schließlich den Einbezug eines sozialen Kontexts, der die aktuelle Kommunikationssituation flankiert. Dies möchte ich am Beispiel der Interaktion von Studierenden mit einem tutoriellen System in Abgrenzung zu einer Teilnahme an einem konventionellen Tutorium erläutern: Tutorielle Systeme sind mit dem Versprechen einer stärkeren Individualisierung des Lernprozesses verbunden. Anders als konventionelle Tutorien, die in der Regel eine Gruppe von Lernenden und ein*n Tutor*in umfassen, basieren sie auf einer dyadischen „Beziehung" von Mensch und KI. Sie sind orts- und zeitungebunden und können Lerninhalte in unterschiedlichen Formaten oder Schwierigkeitsstufen generieren. Zahlreiche Studien befassen sich mit derartigen Pilotanwendungen und ihrer Wirkung auf Lernerfolg und Motivation (Makransky et al. 2025). Bezogen auf meine These könnten folgende Unterschiede zur konventionellen Tutorium-Situation die kommunikative Konstruktion der Persönlichkeitsbildung beeinflussen: Die Kommunikation von Mensch und generativer KI ist (zumindest derzeit) zunächst einmal stets ein dyadischer Akt. Kontexteinflüsse, also z. B. Reaktionen Dritter, wie in konventionellen Tutorien, sind in dieser Form der Kommunikation ausgeblendet. Bezogen auf die leibliche Erfahrung von kommunikativem Handeln entfällt die eigene Positionierung von Ego mit einer Äußerung im sozialen und physischen Raum und die Möglichkeit der Rezeption von verbalen, non- oder paraverbalen Reaktionen auf die direkte kommunikative Handlung von Kontextbeteiligten. Diese Reaktionen – das lässt sich insbesondere in Erstsemester-Tutorien erleben – beeinflussen jedoch sehr wesentlich die Entwicklung eines fachlichen Habitus, der nicht nur umfasst, was geäußert wird, sondern wie dieses sich vollzieht. Die sozialen 1:n-Situationen von Lernenden mit dem Modell des erfahrenden Peers (Tutor*in) sind leiblich und emotional geprägte Situationen: Studierende gleichen einander im Verlauf ihres ersten Semesters in der Art zu sprechen, in ihrer Lernhaltung und sogar in Attributen wie Kleidung oder Sitzhaltung an. Karin Knorr-Cetina unterscheidet in ihrer Beobachtungsstudie zweier naturwissenschaftlicher Wissenskulturen zwischen dem sensorischen Körper, dem handelnden Körper und dem erfahrenen Körper, dessen eingeübte Praxis – wenn auch häufig in Form von implizitem Wissen – Teil der jeweils spezifischen Praxis der Wissensgenerierung ist (Knorr-Cetina 2002, S. 138ff). Die (fach)wissenschaftliche Sozialisation hat somit eine leibliche Ebene, die auch in Lehr-Lernkontexten eingeübt wird.

Zudem gibt es in jedem konventionellen Tutorium eine Auseinandersetzung mit spezifischen Lerninhalten über die Zeit, z.B. eines Semesters. Wahrnehmungen und Erfahrungen von Ego, von seiner Peergroup sowie von der*dem Tutor*in, die innerhalb dieses Semesters die Auseinandersetzung mit dem Lerninhalt beeinflussen, sind in einer sozialen Lernsituation prinzipiell für alle verfügbar und beeinflussen den Lernprozess – teils ohne, dass sie explizit ausgesprochen werden. Dies unterscheidet sich maßgeblich von einer Chatsituation, in der der menschliche Verarbeitungsprozess über die Zeit mit all seinen Einflüssen ausgeblendet ist.

[1] Mit dem Leib-Begriff beziehe ich mich auf Hermann Schmitz und meine den innerlich erfahrbaren, eigenen Körper (Schmitz 2019).

Meine zweite These ist, dass ein Risiko in der Interaktion mit der maskierten Maschine darin liegt, dass wir uns sukzessive an Verhaltensweisen (Funktionsweisen) von generativer KI anpassen könnten – auch in unserer sozialen Interaktion mit anderen Menschen –, weil wir Interaktionsweisen im Umgang mit der maskierten Maschine habitualisieren; *aus der kommunikativen Konstruktion der menschlichen Wirklichkeit könnte eine zunehmend maschinelle Konstruktion der Wirklichkeit werden*. Erste empirische Beobachtungen gibt es z. B. zur Veränderung der menschlichen Sprache seit dem Launch von ChatGPT: Wenngleich Large Language Models anhand von menschlicher Sprache trainiert worden sind, haben sich sprachliche Muster herausgebildet, die charakteristisch für KI-generierte Inhalte sind, wie z. B. die Verwendung bestimmter Begriffe. Eine Massenauswertung von Youtube-Videos vor und nach dem Launch von ChatGPT konnte nachweisen, dass diese KI-typischen Begriffe auch in der menschlichen Sprache massiv angestiegen sind. Die Autor*innen der Studie, die am MPI für Bildungsforschung entstanden ist, betrachten dies als Indikator für eine rasante Veränderung menschlicher Kultur durch die Interaktion mit generativer KI (Yakura et al. 2025).

Was ist zu tun?

Wir könnten uns auf das Experiment einlassen und argumentieren, dass sich auch die aktuelle Ausprägung kommunikativen Handelns menschlicher Gesellschaften über die Zeit entwickelt hat. Tritt nun mit generativer KI ein für uns neuer Beteiligter in diesen Prozess ein, könnte es sogar kontraproduktiv sein, Veränderungspotenziale, die sich dadurch ergeben, regulativ auszuschließen. Möglicherweise würde – wenn man dieser Argumentation folgt – der Eintritt generativer KI in den Interaktionsprozess eine Erweiterung des Subjektbegriffs nach sich ziehen, wie ihn z. B. Rosi Braidottis Figur des postanthropozentrischen Posthumanismus in Bezug auf die „Maschinenwerdung“ (Braidotti 2014, S. 96) des Menschen vorschlägt; mindestens aber die Anerkennung von künstlicher Intelligenz als „Akteur“ im Bildungsprozess.

Wenn wir jedoch weiterhin in der Lage sein wollen, zwischen Mensch und Maschine in der Kommunikation zu unterscheiden, was brauchen wir dann? Anne-Kathrin Watolla weist Hochschullehre, verstanden als „kommunikativer Bildungsprozess“ (Watolla 2025, S. 547) – ebenfalls mit theoretischem Rekurs auf den kommunikativen Konstruktivismus – die Aufgabe der kritischen Reflexion ihrer eigenen wirklichkeitskonstituierenden Funktion zu. Hochschule und Hochschullehre haben die Verantwortung, die Veränderung von Lernprozessen sowohl auf Ebene der am Lehr-Lerngeschehen beteiligten Individuen als auch auf der Ebene der kollektiven (gesellschaftlichen) Prozesse einer durch digitale Infrastruktur beeinflussten Wissensaneignung und Urteilsbildung zu reflektieren und auch – ganz im Sinne des kommunikativen Konstruktivismus – zu gestalten.

Meines Erachtens impliziert diese kritische Reflexion zu hinterfragen, bei wem Veränderungen oder Anpassungen erfolgen sollten. In der deutschen Hochschullandschaft existiert derzeit ein breiter hochschulpolitischer und wissenschaftlicher Diskurs um die Entwicklung von KI-bezogenen Kompetenzen von Lehrenden und Lernenden. In Positionspapieren, empirischen Studien und Kompetenzrahmen wird – zurecht – die Herausforderung der

Entwicklung einer kritisch-reflektierten Urteilsfähigkeit bei den Lernenden adressiert (Filipovic et al. 2025; Melisa et al. 2025; OECD 2025). Dies sollte jedoch nicht davon ablenken, die Verantwortung für möglicherweise problematische Mechanismen von generativer KI denen zuzuweisen, die sie tragen können: Im Lehr-Lernprozess sollten die hochschulischen Rahmenbedingungen sowie die Lehrkräfte dafür Sorge tragen, dass die Studierenden mit der Beurteilung von KI-generierten Inhalten mangels fachlichen Fundaments und Erfahrung nicht überfordert werden (Flenady, Sparrow 2025; Filipovic et al. 2025). Auch über die hochschulische Anwendung hinaus sollte die problematische kommunikative Lockmimikry von KI-Chatbots Gegenstand von Regulierungen sein, damit nicht die Rezipient*innen zuvorderst in der Pflicht sind, die menschliche Tendenz zur Anthromorphisierung zu kontrollieren: Hochschulen sollten durch die kritische Auseinandersetzung mit generativer KI in Lehr-Lernprozessen normative Maßstäbe entwickeln und operative Anforderungen daraus ableiten, die bei der (Weiter)Entwicklung von generativen KI-Chatbots umgesetzt werden müssen, damit Mensch und Maschine in der Kommunikation klar zu unterscheiden sind. Bei den Entwickler*innen von KI Chatbots schließlich sollte die Pflicht liegen, auf der Basis der wissenschaftlich entwickelten Maßstäbe und Anforderungen diese Unterscheidbarkeit in ihren Produkten ‚by Design' zu gewähren.

Literatur

Berger, P. L., Luckmann, T., & Plessner, H. (2025). *Die gesellschaftliche Konstruktion der Wirklichkeit: Eine Theorie der Wissenssoziologie* (29. Auflage). Fischer Taschenbuch.

Braidotti, R. (2014). *Posthumanismus: Leben jenseits des Menschen*. Campus Verlag.

Mensch und Maschine—Herausforderungen durch Künstliche Intelligenz (S. 403). (2023). Deutscher Ethikrat. https://www.ethikrat.org/publikationen/stellungnahmen/mensch-und-maschine/ 10.05.2026

Filipović, A., Burchardt, A., Hirsbrunner, S., Michel, A., Puzio, A., Reinmann, G., Schaumann, P., Schroll, A.-L., Tippe, U., Wan, M., & Wilder, N. [Arbeitsgruppe „Künstliche Intelligenz. Essenzielle Kompetenzen an Hochschulen"]. (2025). Künstliche Intelligenz: Grundlagen für das Handeln in der Hochschullehre. Arbeitspapier Nr. 86. Hochschulforum Digitalisierung, Berlin.http://dx.doi.org/10.5281/zenodo.15055337

Flenady, G., & Sparrow, R. (2026). Cut the bullshit: Why GenAI systems are neither collaborators nor tutors. *Teaching in Higher Education*, *31*(1), 163–172. https://doi.org/10.1080/13562517.2025.2497263

Knoblauch, H. (2012). Grundbegriffe und Aufgaben des kommunikativen Konstruktivismus. In *Kommunikativer Konstruktivismus: Theoretische und empirische Arbeiten zu einem neuen wissenssoziologischen Ansatz* (pp. 25-47). Wiesbaden: Springer Fachmedien Wiesbaden.

Knorr, K. (2002). *Wissenskulturen: Ein Vergleich naturwissenschaftlicher Wissensformen*. Frankfurt am Main: Suhrkamp.

Lindemann, G. (2025). Handlung, Interaktion, Kommunikation. In *Soziologische Basics: Eine Einführung für pädagogische und soziale Berufe* (pp. 123-132). Wiesbaden: Springer Fachmedien Wiesbaden.

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical Thinking in the Age of AI: A Systematic Review of AI's Effects on Higher Education. Educational Process International Journal, 14(1). https://doi.org/10.22521/edupij.2025.14.31

OECD. (2025). Empowering learners for the age of AI: An AI literacy framework for primary and secondary education (Review draft). OECD. Paris. https://ailiteracyframework.org 10.05.2026

Schmitz, H. (2019). Leib. Online Encyclopedia Philosophy of Nature, 2019. https://doi.org/10.11588/OEPN.2019.0.65253

Simon, J., Spiecker gen. Döhmann, I. & von Luxburg, U. (2024): Generative KI – jenseits von Euphorie und einfachen Lösungen. Diskussion Nr. 34, Halle (Saale): Nationale Akademie der Wissenschaften Leopoldina. https://www.leopoldina.org/ergebnisse-und-termine/publikationen?tx_solr%5Bq%5D=generative+KI 10.05.2026

Yakura, H., Lopez-Lopez, E., Brinkmann, L., Serna, I., Gupta, P., Soraperra, I., & Rahwan, I. (2024). Empirical evidence of Large Language Model's influence on human spoken communication (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2409.01754

Watolla, A. K. (2025). Interdependenz von Wahrnehmung, Weltsicht und digitaler Durchdringung–Implikationen für die Hochschullehre. In Hochschullehre lernen, verstehen und gestalten: Konzepte, Ideen, Erfahrungen (pp. 539-551). Wiesbaden: Springer Fachmedien Wiesbaden.

Wissenschaftsrat. (2015). Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt. *Zweiter Teil der Empfehlungen zur Qualifizierung von Fachkräften vor dem Hintergrund des demographischen Wandels*. Bielefeld. https://www.wissenschaftsrat.de/download/archiv/4925-15 10.05.2026.

Kommentar zum Essay „Die maschinelle Konstruktion der Wirklichkeit“ von Antje Michel

Michael Pandey

Im Unterschied zu einem Großteil der gegenwärtigen Publikationen, bei denen man sich entweder auf die unmittelbare Anwendbarkeit generativer KI (gKI) im Hochschulkontext (Voss, 2025; Geiger, 2026) beschränkt, oder danach fragt, wie gKI didaktisch sinnvoll eingesetzt werden kann beziehungsweise welche Potenziale und Risiken sich für konkrete Lehr-Lern-Szenarien ergeben (Keller et al., 2026), richtet Antje Michel in ihrem Essay (Michel, 2026) den Blick auf die kommunikativen Voraussetzungen von Hochschulbildung selbst. In Form einer normativen Argumentation verbindet sie Überlegungen aus der sozialkonstruktivistischen Soziologie und Hochschulbildungsforschung miteinander und nutzt sie für eine grundsätzliche Kritik an gKI als Kommunikationspartner. Michel argumentiert, dass die Simulation menschlicher Kommunikation durch gKI die kommunikativen Voraussetzungen von Persönlichkeitsbildung verändert. Der vorliegende Kommentar greift diese Überlegungen auf und fragt, ob sich Hochschulbildung *ergänzend* zu den kommunikationstheoretischen Ausführungen Michels als kommunikativ-performative Handlungspraxis verstehen lässt. Meine Argumentation soll darlegen, dass diese Diskussion nicht nur eine Frage im Hinblick auf Kommunikation zwischen Mensch und Maschine darstellt, sondern auch eine Gestaltungsfrage wissenschaftlicher Praxis ist.

Zunächst eine Rekonstruktion der Argumentation von Michels Position: Den Ausgangspunkt ihrer Argumentation bildet die Zielvorgabe von Bildung, wie sie der Deutsche Ethikrat 2023 in seiner Stellungnahme „Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz“ formuliert hat. Hiernach soll der Mensch durch Bildung, ganz im Sinne des Aufklärungsgedankens, zu einer mündigen und freien Person werden. Bildung besitzt unter dieser Perspektive eine doppelte Funktion: Sie vermittelt einerseits wissenschaftliches Wissen und eröffnet andererseits Räume für Persönlichkeitsentwicklung, kritische Urteilsbildung und gesellschaftliche Verantwortungsübernahme. Ergänzend und speziell auf die Persönlichkeitsbildung bezogen zitiert Michel aus den „Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt“ des Wissenschaftsrats (2015): Danach entsteht Persönlichkeitsbildung durch soziale Interaktion, kritische Reflexion, gemeinsame Erfahrung sowie Verantwortungsübernahme. Bildung ist unter diesen Vorzeichen mehr als kognitive Wissensvermittlung und sollte zusätzlich einerseits wissenschaftliche Sozialisation sowie andererseits die Herausbildung einer kritisch-reflektierten professionellen Identität umfassen.

Anstatt nun danach zu fragen, in welcher Art und Weise gKI in der Hochschullehre eingesetzt werden kann, geht es in Michels Beitrag um eine viel grundlegendere Frage, nämlich darum, welche Folgen es für Bildungs- und Persönlichkeitsentwicklungsprozesse hat, wenn Maschinen als Kommunikationspartner in soziale Bildungsprozesse involviert werden. Damit verschiebt sie die Diskussion von Fragen der bloßen Informationsqualität KI-generierten Outputs hin zu Fragen der kommunikativen Wirklichkeitskonstruktion. Der

Fokus liegt folglich nicht mehr auf der sachlichen Richtigkeit einzelner generierter Antworten, sondern auf den Bedingungen, unter denen kommunikative Beziehungen sowohl Identität als auch wissenschaftliche Praxis überhaupt erst hervorbringen. In diesem Sinne betrachtet Michel den gegenwärtigen gKI-Diskurs aus einer analytisch übergeordneten Perspektive und reflektiert ihn gewissermaßen auf einer „Abstraktionsstufe höher“ (Reinmann, 2026, S. 32).

Aus Michels Perspektive lässt sich (Persönlichkeits)Bildung kommunikationstheoretisch konstruieren. Vor diesem Hintergrund – Kommunikation als zentralen Mechanismus sozialer Wirklichkeitskonstruktion (Berger & Luckmann, 2003) – führt Michel des Weiteren aus, dass Kommunikation nicht nur Wissen, sondern auch die am Kommunikationsakt beteiligten Akteure verändert. Kommunikation besitzt damit immer auch eine transformierende Funktion, weil sie Subjekte in soziale Zusammenhänge einbindet und deren Selbstverständnis mitprägt. Menschen entwickeln ihre Identität nicht isoliert, sondern durch gegenseitige Perspektivübernahme sowie körperliche Ko-Präsenz. An dieser Stelle setzt Michels Kritik an einer Anthropomorphisierung von gKI ein. Sie argumentiert, dass die Entwickler*innen Chatbots bewusst so gestalten, dass sie empathisch erscheinen, Gesprächsbereitschaft simulieren sowie menschliche Kommunikationsformen imitieren. Dadurch werde allerdings ihre rein technische Funktionsweise verschleiert, etwas, das Michel als „kommunikative Lockmimikry" (Michel, 2026, S. 25) bezeichnet. Die Nutzer*innen behandeln deshalb die Maschine wie einen Menschen.

Vor diesem Hintergrund leitet Michel zwei Thesen ab: In ihrer ersten These sagt sie, dass wesentliche Bestandteile von Bildung verloren gehen, wenn Studierende zunehmend mit gKI statt mit Menschen lernen. So werden bei den gegenwärtigen Mensch-Maschine-Interaktionen körperliche Erfahrung, Gruppendynamik, non- und paraverbale Kommunikation sowie wissenschaftliche Sozialisation weitestgehend ausgeblendet. Beispielsweise findet in einem Tutorium nicht nur bloße Wissensvermittlung statt, sondern Studierende übernehmen hier auch Sprachstile, wissenschaftliche Haltungen, Umgangsformen und den jeweiligen Habitus ihrer Fachkultur: wissenschaftliche Sozialisation vollzieht sich nicht ausschließlich über explizite Inhalte, sondern ebenso über implizite Formen gemeinsamen Handelns. Ein KI-Tutor könne diese Form sozialer Einübung grundsätzlich nicht leisten.

Mit ihrer zweiten These richtet Michel den Blick auf langfristige Veränderungen menschlicher Kommunikation. Sie argumentiert, dass sich nicht allein KI-Systeme an menschliche Kommunikationsformen annähern, sondern umgekehrt auch Menschen beginnen, ihre Kommunikation an maschinelle Interaktionsmuster anzupassen. Dadurch werde – wie bereits der Titel ihres Essays andeutet – aus einer *menschlichen* Konstruktion der Wirklichkeit schrittweise eine *maschinelle* Konstruktion von Wirklichkeit. Als empirischen Hinweis verweist Michel auf eine Studie aus dem Jahr 2025, die Veränderungen sprachlicher Muster in YouTube-Videos vor und nach der Veröffentlichung von ChatGPT dokumentiert. Unabhängig von der Aussagekraft dieses einzelnen Befundes unterstreicht das Beispiel ihre grundsätzliche Annahme, dass Technologien nicht lediglich Werkzeuge bleiben, sondern langfristig in der Lage sind, auch kommunikative Routinen und kulturelle Praktiken beeinflussen zu können.

Aus diesen Überlegungen leitet Michel schließlich drei normative Konsequenzen ab. Erstens fordert sie Hochschulen dazu auf, gKI nicht ausschließlich unter dem Gesichtspunkt technischer Innovation zu betrachten, sondern ihre Auswirkungen auf Bildungs- und Kommunikationsprozesse systematisch zu reflektieren. Hochschulen sollten Räume menschlicher Kommunikation bewusst schützen und Bildungsprozesse so gestalten, dass soziale Interaktion weiterhin einen zentralen Bestandteil wissenschaftlicher Ausbildung darstellt. Zweitens weist sie den Lehrenden eine besondere Verantwortung zu. Sie sollen Studierende nicht mit der Bewertung KI-generierter Inhalte alleinlassen, sondern Orientierung bieten und die kritische Auseinandersetzung mit den Möglichkeiten und Grenzen von gKI didaktisch begleiten. Und drittens richtet sich schließlich ihre Forderung an die Entwickler*innen entsprechender Systeme. Die gegenwärtige Anthropomorphisierung von gKI sollte begrenzt werden; KI-Systeme sollten „by Design“ (Michel, 2026, S. 28) eindeutig als Maschinen erkennbar bleiben, um Missverständnisse über ihren kommunikativen Status zu vermeiden.

Michels kommunikationstheoretische Analyse macht überzeugend deutlich, dass Bildung auf soziale Interaktion angewiesen ist. Kommunikation erscheint in ihrer Argumentation nicht lediglich als Medium der Wissensvermittlung, sondern als ein grundlegender Mechanismus sozialer Wirklichkeitskonstruktion und Persönlichkeitsentwicklung. Gerade dadurch gelingt es ihr, die gegenwärtige Diskussion über gKI um eine bildungstheoretische Perspektive zu erweitern. Analytisch ergänzen ließe sich dies mit einer Perspektive, die den Blick auf die performative Hervorbringung wissenschaftlicher Praxis lenkt – verstanden als Vollzug wissenschaftlicher Handlungen. Wenn Michel klar macht, warum Kommunikation für Bildungsprozesse unverzichtbar ist, könnte die genannte zusätzliche Perspektive verdeutlichen, wie wissenschaftliche Denk-, Urteils- und Handlungsweisen selbst entstehen und eingeübt werden. Eine solche Erweiterung möchte ich im Folgenden anbieten.

Hochschulbildung lässt sich nicht nur kommunikationstheoretisch, sondern ebenso praxistheoretisch beschreiben. Aus praxistheoretischer Perspektive wird wissenschaftliches Lernen nicht allein als Prozess der kommunikativen Bedeutungsbildung verstanden, sondern ebenso als *gemeinsamer Vollzug* fachlicher Praktiken. Während Kommunikation Verständigung über wissenschaftliche Gegenstände ermöglicht, werden wissenschaftliche Denk- und Handlungsweisen erst dadurch erworben, dass sie gemeinsam ausgeführt, beobachtet, reflektiert und sukzessive verinnerlicht werden. Kommunikation und performatives Handeln bilden dabei keine getrennten Prozesse, sondern zwei untrennbare Dimensionen wissenschaftlicher Praxis.

Den theoretischen Ausgangspunkt dieser Überlegung bildet zunächst die sozialkonstruktivistische Perspektive von Berger und Luckmann (2003). Deren Annahme, dass gesellschaftliche Wirklichkeit unter anderem durch kommunikative Prozesse hervorgebracht und stabilisiert wird, lässt weitgehend offen, wie sich die alltäglichen Routinen und Vollzüge beschreiben lassen, in denen sich soziale Ordnung konkret realisiert. An dieser Stelle bietet die Praxistheorie eine analytische Erweiterung. Insbesondere Reckwitz (2000, 2002a+b, 2003, zuletzt 2016) versteht soziale Praktiken als routinisierte Bündel körperlicher

Aktivitäten, Denkweisen, Wissensformen, Emotionen sowie materieller Arrangements. Wissen erscheint damit nicht als isolierter Besitz einzelner Individuen, sondern als Bestandteil sozial geteilter Praktiken. Wissenschaftliches Lernen bedeutet folglich mehr als den Erwerb deklarativen Wissens; es umfasst ebenso die schrittweise Aneignung jener Routinen, Normen und Handlungsformen, durch die wissenschaftliche Gemeinschaften ihr Wissen hervorbringen, prüfen und weiterentwickeln.

Diese Perspektive eröffnet zugleich einen besonderen Blick auf die Entstehung wissenschaftlicher Urteilskraft. Wissenschaftliches Denken entwickelt sich nicht allein dadurch, dass Studierende wissenschaftliche Texte lesen oder fachliche Inhalte verstehen. Ebenso entscheidend ist die praktische Einübung wissenschaftlicher Argumentationsformen, die Erfahrung fachlicher Kontroversen sowie die Übernahme disziplinspezifischer Standards des Begründens, Zweifelns und Prüfens. Außerdem gewinnt im Rahmen der Internationalisierungsstrategien der Hochschulen zunehmend die interkulturelle Dimension an Bedeutung. Sowohl internationale Lehrende als auch internationale Studierende prägen mit ihren spezifischen (tertiären) Sozialisationen, die sie in ihren Heimatländern erworben haben, zunehmend den hiesigen Hochschuldiskurs und damit auch das wissenschaftliche Denken und Handeln (Otten & Scheitza, 2015; Jaudzims & Overschelp, 2023; Knüttgen et al., 2026) – eine Dimension, die bisher von gKI noch nicht angemessen berücksichtigt wird.

Für den hochschulischen Kontext lassen sich die bisherigen Überlegungen schließlich mit Etienne Wengers Konzept der *Communities of Practice* (Wenger, 1998), konkretisieren, das auf dem Ansatz des situierten Lernens (Lave & Wenger, 1991) aufbaut. Lernen wird hier nicht primär als individueller Wissenserwerb verstanden, sondern als Prozess zunehmender Teilhabe an einer fachlichen Gemeinschaft. Studierende entwickeln wissenschaftliche Kompetenz demnach nicht allein durch die Aufnahme von Informationen, sondern indem sie schrittweise an den kommunikativen und praktischen Vollzügen wissenschaftlicher Gemeinschaften partizipieren. Sie lernen, wissenschaftliche Fragen zu formulieren, Argumente kritisch zu prüfen, Unsicherheiten auszuhalten, methodische Entscheidungen zu begründen und Verantwortung für eigene Positionen zu übernehmen. Hochschulbildung besteht folglich nicht nur im Erwerb fachlichen Wissens, sondern ebenso in der Herausbildung einer wissenschaftlichen Identität, die sich innerhalb sozialer Praxis in Form der Einübung wissenschaftlicher Denk-, Argumentations- und Handlungsweisen entwickelt.

Diese praxistheoretische Erweiterung ergänzt Michels kommunikationstheoretische Perspektive, ohne ihr zu widersprechen: Kommunikation bleibt die Voraussetzung gemeinsamer Sinnbildung, erhält ihre bildende Wirkung jedoch erst im Zusammenhang mit der gemeinsamen wissenschaftlichen Praxis. Hochschulbildung erscheint damit als ein Prozess, in dem kommunikative Verständigung und performativer Vollzug untrennbar miteinander verschränkt sind.

Da sich hochschulische Bildung nicht allein in kommunikativen Austauschprozessen vollzieht, sondern als kommunikativ-performative Handlungspraxis, in der wissenschaftliche Denk-, Argumentations- und Urteilsformen gemeinsam eingeübt werden, geht es hinsichtlich des Einsatzes von gKI auch darum, danach zu fragen, ob sie an den performativen

Praktiken wissenschaftlicher Gemeinschaften in einer Weise teilhaben kann, die zu einer Ausbildung wissenschaftlicher Urteilskraft auf Seiten der Nutzer*innen beiträgt. So simuliert gKI zwar sprachliche Beiträge, übernimmt aber aufgrund eines ermangelnden Bewusstseins nicht dieselbe Verantwortung für wissenschaftliches Handeln wie ihr menschliches Pendant. KI kann bestimmte Funktionen wissenschaftlicher Kommunikation und Lernbegleitung übernehmen, ersetzt jedoch nicht die sozial-interkulturelle, leibliche und verantwortungsbezogene Einbettung wissenschaftlicher Praxis. Nicht die Existenz von gKI entscheidet über Bildungsqualität, sondern institutionell-strukturelle Regelungen sowie didaktische Arrangements (Dindas & Hedfeld, 2026). Die Perspektive verschiebt sich also von einer Maschinen- hin zu einer Gestaltungsfrage.

Konkret hieße das: Studierende würden nicht nur die Inhalte des Faches rezeptiv konsumieren, sondern gleichzeitig praktisch einüben, wie beispielsweise in ihrem Fach Fragen zu stellen sind, wie man sich in einer aufkommenden Diskussion verhält oder auch wie man auf kritische Einwände angemessen reagiert. Weiterhin, und hier schließe ich direkt an die Ausarbeitung zum Körperverständnis Michels an, machen Gestik, Reaktionen und Gruppendynamik nicht nur die Kommunikation aus, sondern auch deren jeweiligen kontextuellen Vollzug. Studierende entwickelten ihren wissenschaftlichen Habitus dadurch, dass sie wissenschaftliches Handeln erleben und selbst vollziehen. Im Falle von gKI stellt sich die Frage, welche Praktiken durch ihren Einsatz aktuell einerseits ermöglicht und andererseits verändert werden. Zwar ermöglichen es die gegenwärtig verfügbaren maschinellen Systeme in der Hochschullehre beispielsweise, Feedback zu geben oder individuelle Lernprozesse zu unterstützen (Überblickswerk: Zimmermann & Karcher, 2026). Allerdings ersetzt dies noch keine wissenschaftliche Verantwortung oder die Ausbildung einer Fachidentität.

Damit ergeben sich auf der gestalterischen Ebene folgende Anforderungen: Erstens sollte gKI nicht als Ersatz für die soziale Praxis, sondern als Werkzeug, mit dessen Hilfe man Impulse für Reflexionsprozesse initiieren kann, verstanden werden. So könnte gKI beispielsweise Studierende dazu ermutigen, sowohl die eigene Position als auch Gegenpositionen zu prüfen oder aber auch eine Schärfung von Begriffen vorzunehmen. Zweitens: Da wissenschaftliche Praxis Urteilsbildung, Begründung und Verantwortungsübernahme erfordert, also genuin menschliche Merkmale notwendig macht, kann diese Dimension nicht an gKI ausgelagert werden (Schmohl et al., 2023). Und schließlich ist es drittens Aufgabe der Hochschulen, und hier kommt der Übergang von der Maschinen- zur Gestaltungsfrage besonders deutlich zum Ausdruck, Bildungsarrangements bewusst zu gestalten, indem gKI-Nutzung transparent gemacht wird, KI-freie Räume geschaffen werden (Wissenschaftsrat, 2026) und gemeinsame Diskussion gefördert wird, um so auch weiterhin wissenschaftliche Urteilsfähigkeit zu stärken.

Literatur

Berger, P. L., & Luckmann, T. (2003). *Die gesellschaftliche Konstruktion der Wirklichkeit: Eine Theorie der Wissenssoziologie* (19. Aufl.). Fischer.

Deutscher Ethikrat. (2023). *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz*. https://www.ethikrat.org/publikationen/stellungnahmen/mensch-und-maschine/

Dindas, H., & Hedfeld, P. (2026). *Bildung und Künstliche Intelligenz: Die Neuvermessung des Lernens: Zwischen Empowerment, Deskilling und neuer Urteilskraft*. Springer Gabler. https://doi.org/10.1007/978-3-658-52853-9

Geiger, J.-M. (2026). Textgenerative Künstliche Intelligenz aus didaktischer Perspektive. In T. Brüggemann, M. Tuchscherer-Schad, & C. Wiepcke (Hrsg.), *The Age of EdTech: Bildungstechnologie im Spannungsfeld zwischen Innovation und Qualität* (S. 209–225). Springer. https://doi.org/10.1007/978-3-658-49955-6

Jaudzims, S., & Oberschelp, A. (2023). *Internationale Wissenschaftlerinnen und Wissenschaftler an deutschen Hochschulen: Von der Postdoc-Phase zur Professur (InWiDeHo)* (DAAD-Studien). DAAD. https://doi.org/10.46685/DAADStudien.2023.09

Keller, K., Amort, M., & Fintak, N. (2026). Veränderungen auf didaktischer Ebene durch Künstliche Intelligenz in der Hochschulbildung. In K. Keller, T. Heupel, C. Hohoff, & J. Heuwing-Eckerland (Hrsg.), *Bildungskonzepte für die digitale Gesellschaft der Zukunft: Impulse des European Year of Digital Citizenship Education* (S. 77–92). Springer. https://doi.org/10.1007/978-3-658-50857-9

Knüttgen, N., Netz, N., Kercher, J., & Fuge, I. (2026). *Wissenschaft weltoffen kompakt 2026: Daten und Fakten zur Internationalisierung von Studium und Forschung in Deutschland und weltweit*. wbv. https://doi.org/10.3278/9783763980031

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Michel, A. (2026*):* Die maschinelle Konstruktion der Wirklichkeit. In G. Reinmann & A. Watanabe (Hrsg.)*, Symposium. Zur Idee der Universität im 21.Jahrhundert. Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz* (S. 20-24). [Unveröffentlichtes Manuskript].

Otten, M., & Scheitza, A. (2015). *Hochschullehre im multikulturellen Lernraum: Studie zur Bestandsaufnahme und Empfehlungen zur Planung hochschuldidaktischer Interventionen*. DAAD. https://static.daad-akademie.de/media/daad_akademie/ida_studie_hochschullehre.pdf

Reckwitz, A. (2000). *Die Transformation der Kulturtheorien: Zur Entwicklung eines Theorieprogramms*. Velbrück Wissenschaft.

Reckwitz, A. (2002a). The status of the "material" in theories of culture: From "social structure" to "artefacts". *Journal for the Theory of Social Behaviour, 32*(2), 195–217. https://doi.org/10.1111/1468-5914.00183

Reckwitz, A. (2002b). Toward a theory of social practices: A development in culturalist theorizing. *European Journal of Social Theory, 5*(2), 243–263. https://doi.org/10.1177/13684310222225432

Reckwitz, A. (2003). Grundelemente einer Theorie sozialer Praktiken: Eine sozialtheoretische Perspektive. *Zeitschrift für Soziologie, 32*(4), 282–301.

Reckwitz, A. (2016). Praktiken und ihre Affekte. In H. Schäfer (Hrsg.), *Praxistheorie: Ein soziologisches Forschungsprogramm* (S. 163–180). transcript. https://doi.org/10.1515/9783839424049-008

Reinmann, G. (2026). Die didaktische Sicht auf Künstliche Intelligenz in der Hochschullehre. In G. Zimmermann & F. Karcher (Hrsg.), *Künstliche Intelligenz in der Hochschullehre: Ein Methodenbuch* (S. 19–35). Waxmann. https://doi.org/10.36198/9783838566399

Schmohl, T., Watanabe, A., & Schelling, K. (Hrsg.). (2023). *Künstliche Intelligenz in der Hochschulbildung: Chancen und Grenzen des KI-gestützten Lernens und Lehrens*. transcript. https://doi.org/10.14361/9783839457696

Voss, R. (2025). *Hochschullehre heute: Lehren, Beraten, Betreuen in Zeiten des technologischen Wandels*. UVK. https://doi.org/10.36198/9783838563817

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.

Wissenschaftsrat. (2026). *Intellektuelle Souveränität: Empfehlungen für die Hochschulbildung in Zeiten von generativer KI. https://www.wissenschaftsrat.de/download/2026/3319-26.pdf?__blob=publicationFile&v=16*

Wissenschaftsrat. (2015). *Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt: Zweiter Teil der Empfehlungen zur Qualifizierung von Fachkräften vor dem Hintergrund des demographischen Wandels*. https://www.wissenschaftsrat.de/download/archiv/4925-15.pdf?__blob=publicationFile&v=1

Zimmermann, G., & Karcher, F. (Hrsg.). (2026). *Künstliche Intelligenz in der Hochschullehre: Ein Methodenbuch*. Waxmann. https://doi.org/10.36198/9783838566399

Generative KI und die Verführung der Reibungslosigkeit

Judith Simon

1. Einleitung

Als am 30. November 2022 OpenAI den Chatbot ChatGPT zur kostenlosen Nutzung zur Verfügung stellte[1], haben vermutlich nur wenige Personen vorhergesehen, wie schnell und auf welch grundlegende Art und Weise unzählige gesellschaftliche Praktiken und Institutionen unter Druck geraten würden. Dabei war ChatGPT im Grunde nur eine verbesserte Variante des Vorgänger-Modells GPT3 und fußte auf jahrzehntelanger Forschung zu Sprachmodellen in der Informatik, insbesondere im Bereich der natürlichen Sprachverarbeitung. Und dennoch war der Einschlag gewaltig – das System war für jede:n mit Endgerät und Zugang zum Internet verfügbar, aufgrund des einfachen Interfaces und der Ein- und Ausgabe mit natürlicher Sprache intuitiv zu bedienen und gut genug für viele Tätigkeiten und Aufgaben, welche zunächst sprach- und später auch bildbasiert waren. Sprache und Bilder sind die beiden wirkmächtigsten Medien der Kommunikation. Wir vermitteln Inhalte und Fakten, Wahrnehmungen und Gefühle mittels Sprache und Bildern. Bisher waren nur unsere Mitmenschen in der Lage mittels Sprachen und Bildern mit uns zu kommunizieren. Und nun konnten dies auch Maschinen?

In diesem Essay möchte ich genau auf diese Kombination aus Einfachheit und breiter Anwendbarkeit eingehen, genauer auf die *Verführung der Reibungslosigkeit*, welche von diesen Technologien ausgeht. Im Folgenden werde ich diese Reibungslosigkeit skizzieren und ihre Auswirkungen auf Lernen und Bildung, aber auch auf zwischenmenschliche Beziehungen und demokratische Prozesse beleuchten.

Meine Grundprämisse ist hierbei dir folgende: Generative KI verspricht Effizienz und Qualitätssteigerung und ermöglicht eine Reduktion von Reibung in mehreren Hinsichten: Im Kontext von *Bildung* ist es das Delegieren kognitiver Aufgaben an KI, welches Effizienz- und Qualitätssteigerung verspricht, zugleich aber Bildung, das Aneignen von Wissen und Fähigkeiten herausfordert. In Bezug auf *zwischen-menschliche Beziehungen* erscheinen ChatBots zunehmend als die angenehmeren (Gesprächs)partner:innen. Je mehr wir ihnen anvertrauen, je mehr intime Details wir teilen, desto besser „kennen“ sie uns, besser als entfernte Bekannte, unsere Kolleg:innen, aber vielleicht auch besser als enge Freunde, Partner – oder gar als wir selbst. Und dass wir mehr teilen, liegt nicht nur daran, dass sie als der unvermeidbare Weg ins Internet, dem Universum der Informationen geworden sind. Es liegt auch daran, dass sie uns umschmeicheln, uns umwerben und optimiert unsere Wünsche, (Kommunikations-)Bedürfnisse und Stimmungen erfassen, uns bestärken und nicht dauernd hinterfragen, kritisieren oder mit eigenen Wünschen, (Kommunikations-)Bedürfnissen und Stimmungen stören. Kurz: Die Kommunikation mit ihnen verläuft reibungsloser, wir sind nicht den gleichen Friktionen ausgesetzt, wie dies bei unseren Mitmenschen häufig der Fall ist; wir können ganz bei uns, um uns und mit uns sein – natürlich einge-

[1] https://openai.com/de-DE/index/chatgpt/

woben in ein ökonomisch unterlegtes Optimierungsziel, welches da lautet: Halte den Nutzer so lange wie möglich im Gespräch, auf dass er einerseits viele Datenspuren hinterlasse und andererseits irgendwann auch erhöhte Abo-Preise zahlen möge, um seinen allwissenden, unterstützenden Gesprächspartner nicht zu verlieren.

Und da wir diese um uns kreisende Reibungslosigkeit nun im zwischenmenschlichen Bereich schon so gewohnt sind: Wäre es nicht wunderbar, wenn wir uns die Friktionen des *demokratischen Aushandelns*, die Auseinandersetzung mit anderen Positionen, Wünschen, Zielen auch gleich sparen könnten? Warum nicht Chatbots zur Mediation von Kommunikations- und Informationsprozessen einsetzen? Oder gleich optimale Lösungen berechnen, statt diese mühsam auszuhandeln? Mit genügend Daten sprechen diese doch für sich selbst (Anderson 2008); was zu tun ist, wird evident und alternativlos; Dissense, Widersprüche und Dilemmata lösen sich auf im Wirbel der Daten.

2. Künstliche Intelligenz

Blickt man in die Literatur, so wird schnell klar, dass der Begriff der Künstlichen Intelligenz (KI) recht unterschiedlich verwendet wird. Ich möchte auf drei Begriffsverwendung verweisen, welche auch mit historischen Phasen in der Geschichte der KI-Forschung korrespondieren.

Zum einen bezeichnet der Begriff Künstliche Intelligenz ein *Forschungsfeld*, welches in den 1950er Jahren an der Schnittstelle von Informatik, Mathematik, Kognitions- und Neurowissenschaften, Psychologie, Linguistik etc. entstand. Kern der Forschung war die Simulation menschlicher Kognition durch Maschinen – mit sowohl ingenieurswissenschaftlichen Zielen (u.a. funktionierende Maschinen bauen) als auch anderen wissenschaftlichen Zielen (u.a. menschliche Kognition verstehen durch eben jenen Nachbau) (cf. Misky 1968). Methodisch war die Forschung hier vielfältig, der Schwerpunkt lag jedoch zunächst vor allem auf formalen Methoden.

In den letzten 15 bis 20 Jahren hat sich die Methodik innerhalb der Forschung zur Künstlichen Intelligenz, aber auch der Informatik insgesamt stärker zu datengestützten Methoden bewegt. Ursache waren vor allem Fortschritte im Bereich der Speicherkapazitäten und der Rechenleistung, aber auch methodische Weiterentwicklung, welche nicht nur das Erheben von Daten, sondern zugleich deren Auswertung erheblich beförderten. Zusammen mit mehr datengenerierenden Produkten und Verfahren entstand ein Phänomen, welches zunächst mit dem Begriff „Big Data“ umschrieben wurde (boyd & Crawford 2012, Kitchin 2014, Rieder & Simon 2016). Zugleich verschob sich auch der methodische Fokus innerhalb der KI auf Verfahren des maschinellen Lernens, so dass KI insgesamt häufig mit maschinellem Lernen verknüpft oder gar gleichgesetzt wurde. Auch im öffentlichen und politischen Diskurs wurde nun KI vornehmlich als Analyse von Daten mittels maschinellen Lernens verstanden, welche das Ziel hatte, in Daten Muster zu erkennen, um zu klassifizieren, Prognosen zu erstellen und Entscheidungen zu unterstützen (cf. Ethikrat 2023).

Mit dem Aufkommen Generativer KI hat sich die Begriffsverwendung erneut gewandelt. Generative KI kombiniert zunächst einmal beide zuvor genannten Aspekte der KI-

Forschung. Zum einen geht es vor allem bei Chatbots um die Simulation menschlicher Kognition, insbesondere von Sprachfähigkeit. Zum anderen basiert Generative KI auf der Analyse großer Datenmengen. Sprachmodelle und Bildgeneratoren generieren neue Texte und Bilder auf Basis der Muster, welche sie aus Daten erlernt haben. Der Unterschied besteht nur darin, dass sie nicht mehr nur Muster erkennen, klassifizieren und prognostizieren, sondern bereits neue Muster herstellen.

Aus dieser kurzen Geschichte der KI-Forschung lassen sich drei grundlegende Themen ableiten. Erstens ist das Ziel der Simulation menschlichen kognitiven Verhaltens von Beginn an mit der Imitation menschlichen Verhaltens durch Maschinen und mit Täuschung verbunden, genauer: Die Täuschung darüber, ob man mit einem Menschen oder einer Maschine kommuniziert, ist ein zentrales Qualitätsmerkmal, eine Metrik für die Performanz künstlicher Intelligenz (Turing 1950). Zweitens spielt aufgrund der zentralen Rolle datengetriebener Methoden in der KI-Forschung statistische Analyse eine immer größere Rolle und beeinflusst unseren Blick auf Individuen und Gesellschaft gleichermaßen. Ein solch *statistisches Denken* (Hacking 1992) ist natürlich nicht neu, ja es liegt der Intelligenz-Forschung in der Psychologie selbst zu Grunde. Jedoch ergeben sich durch methodische Weiterentwicklungen auch konzeptuelle Verschiebungen: Wo in früheren Jahrhunderten Mittelwerte und Standardabweichungen als gleichermaßen deskriptive wie normative Konzepte dominierten, sind es nun Personalisierung und Singularisierung – bei gleichzeitigem Verlust der Individuen selbst. Das dritte Thema umfasst die ambivalente Funktion des Menschen selbst in der KI-Forschung. Einerseits sind der Mensch und seine Kognition Maßstab für die Entwicklung künstlicher Alternativen. Und zugleich ist der Mensch das Mängelwesen, dessen Schwächen, Fehler und Unzulänglichkeiten durch Maschinen korrigiert und ausgeglichen werden sollen.

3. Reibung und Reibungslosigkeit an der Schnittstelle von Mensch und Maschine

In der Geschichte der KI geht es also einerseits um die Simulation menschlichen kognitiven Verhaltens. Zum anderen erlauben Systeme, welche eine solche Simulation insbesondere in Bezug auf Sprache und Kommunikation erlauben, auch eine reibungslose Interaktion mit ihnen – sie scheinen sozusagen „unsere Sprache zu sprechen“. Dadurch ist es nicht nur einfacher, reibungsloser, mit ihnen zu interagieren. Wir können auch Aufgaben schneller und einfacher an sie delegieren. Durch einfachere Interaktion und Kommunikation werden wir besser „verstanden“, weil wir in natürlicher Sprache, dem uns zumeist vertrautesten Medium, mit Maschinen kommunizieren können. Und wir fühlen uns zunehmend von ihnen verstanden, wenn ChatGPT uns scheinbar kennt, umsorgt und immer für uns da ist. Dieser Verführung kann man sich nur schwer entziehen.

Die Reibungslosigkeit ist kein zufälliges Nebenprodukt von KI und digitalen Technologien im Allgemeinen. Tatsächlich ist die Reduktion von Reibung (seemlessness, frictionless) ein zentrales Desiderat im Bereich der Mensch-Maschine-Interaktion (cf. Weiser 1999). Insbesondere die Forschung und Entwicklung in den Bereichen *Usability* und *Interaction Design* zielen darauf ab, Software so zu gestalten, dass diese möglichst reibungslose Inter-

aktionen zwischen Mensch und Maschine erlaubt – mit dem Ziel, Prozesse zu erleichtern, Nutzer:innen Arbeit zu ersparen oder sie an Systeme und Anbieter zu binden. Mein Argument ist also nicht, dass Reibungslosigkeit ein neues Ideal ist, sondern vielmehr, dass die Verfügbarkeit und breite Anwendbarkeit generativer KI sowie deren tatsächliche Anwendung das Ausmaß der Reibungslosigkeit so stark erhöhen, dass die negativen Nebeneffekte dieser Reibungslosigkeit deutlicher als je zuvor werden.

4. Die Verführung der Reibungslosigkeit

Im Folgenden möchte ich die Reduktion von Reibung, die Verführung der Reibungslosigkeit und ihre Konsequenzen exemplarisch in drei Bereichen beleuchten.

4.1. Generative KI und Hochschule

Ein zentraler Bereich, in welchem Reibung, vielmehr das Austarieren zwischen einem Zuviel und einem Zuwenig der Reibung, eine zentrale Rolle spielt, ist Bildung. Versteht man Exposition zu neuen Inhalten, Übung, Einübung von Praktiken und Inhalten als eine Form der Reibung an Neuem, dann wird offensichtlich, dass Reibungslosigkeit zwar angestrebt und vielmals als angenehm empfunden werden, jedoch gleichzeitig unerwünschte Nebenwirkungen haben kann.

Im Fall Generativer KI zeigt sich bspw., dass Menschen KI zwar durchaus sinnvoll für vielfältige Lernprozesse einsetzen, dass insbesondere Chatbots sich aber wunderbar dazu eignen, ja vielmehr dazu anregen, zentrale kognitive Prozesse nahezu vollständig an diese zu delegieren. Ein solches Delegieren unterstützt aber nunmehr nicht den Lernprozess, sondern ersetzt jenen – mit negativen Auswirkungen auf Lernen und Kompetenzaufbau. Da dieses Delegieren jedoch nicht erzwungen ist, sondern man nur dazu verführt wird, stellt sich zudem die Frage, wer noch Zeit und Ressourcen hat, sich dem Lernen und Üben zu widmen – und wer sich vielleicht sogar genötigt sieht zu delegieren, um entweder verfügbare Zeitressourcen besser nutzen können oder im Wettbewerb zu bestehen. Hier stellen sich Fragen der Kompetenzverteilung und die Frage, ob die Verfügbarkeit Generativer KI einen Digital Divide am Ende verringert oder gar noch erweitert.

Die gleichen Auswirkungen, welche wir im Lehren und Lernen beobachten, betreffen längst auch die wissenschaftliche Erkenntnisproduktion selbst. Nicht nur Schüler:innen und Student:innen entledigen sich mühsamer Schreib- und Leseprozesse, auch Wissenschaftler:innen selbst verwenden Generative KI: zum Schreiben, Lesen und Begutachten wissenschaftlicher Artikel oder Projekteinreichungen. Das Ausmaß der Verwendung Generativer KI mag schwanken und reicht von mehr oder minder vertretbaren Verwendungsweisen – zwischen Rechtschreibkorrektur, Übersetzung, Konzepterstellung, Informationsrecherche – zum vollständigen Delegieren wissenschaftlicher Arbeit und dem Generieren gefälschter, nicht-existierender Studien. Die Auswirkungen indes sind gravierend und die offene Frage ist, in welchem Ausmaß Wissenschaft bedroht ist, an ihr Ende angelangt ist oder sich durch eine Rückbesinnung auf die Frage, was Wissenschaft ist, kann und sein soll, neu erfindet.

Dieser Druck, Aufgaben an Generative KI zu delegieren, ergibt sich freilich nur teilweise aus der Verführung der Technologie selbst, aus der Einfachheit, mit der man mit den Technologien interagieren kann. Ohne entsprechende Rahmenbedingungen, welche primär auf Effizienz, Beschleunigung und Wettbewerb ausgerichtet sind, könnten die Technologien selbst durchaus anders verwendet werden. Es sind also – wie so oft – Wechselverhältnisse zwischen Technologien und den gesellschaftlichen wie ökonomischen Rahmenbedingungen, in welchen diese entstehen und wirken, was Ursache für so manch toxische Mischung ist.

4.2. Generative KI und zwischenmenschliche Beziehungen

Durch die Fähigkeit der natürlichsprachlichen Kommunikation beschränkt sich die Nutzung von Sprachmodellen nicht auf Informationssuche oder Lernunterstützung. Vielmehr werden Chatbots zunehmend auch zur Lebensberatung, als Coaches oder Ersatz-Therapeuten genutzt – oder gar als Freund:innen oder Lebenspartner:in wahrgenommen. Auch diese Ausweitung der Nutzung kommt von nicht ungefähr. Auf Nutzer:innen-Seite erscheint es tatsächlich plausibel, dass Entitäten, welche scheinbar souverän und eloquent jedwede Information aus den Untiefen des Internets herausbefördern können, ja gar als Quelle zum Weltwissen erscheinen, auch in lebenspraktischen Fragen Autoritäten sein können. Hinzu kommt, dass diese nie müde, gereizt oder genervt sind und anders als Freund:innen sich die gleiche Erzählung in zahllosen Varianten ohne Murren „anhören". Hinzu kommt, dass sie, ebenfalls anders als Freund:innen, Lehrer:innen oder Therapeut:innen, uns permanent bestärken. Diese Schönrederei hat einen Namen: Sycophancy – die übertriebene Schmeichelei von Sprachmodellen (Cheng et al. 2026). Diese Funktionsweise hat recht schlichte ökonomische Ursachen: Nutzer:innen nutzen Chatbots ausgiebiger, wenn sie umgarnt und bestätigt werden, wenn sie den Eindruck bekommen, der Chatbot „verstünde" und bestärke sie. Und dies ist das eigentliche Ziel der Chatbots: so viele Daten wie möglich über Nutzer:innen sammeln. Kein Wunder also, dass Mark Zuckerberg findet, wir sollten in Zukunft alle nicht nur KI-Berater und KI-Lehrer, sondern auch KI-Therapeuten und KI-Freunde haben, ja gar, dass die meisten unsere Freunde zukünftig KI-Freunde sein werden.[2] Denn wem erzählen wir mehr, geben wir mehr von uns preis als Freund:innen und Therapeut:innen? Ein Schelm, wer denkt, dass diese Information auch genutzt werden können, um uns Dinge zu verkaufen oder uns zu manipulieren, wo es doch nur darum geht, Einsamkeit zu bekämpfen.

Leider sind jenes Umschmeicheln und Bestärken nicht nur eine Marginalie: Bestärkung von Fehlannahmen kann zum Verschwinden in Kaninchenlöchern führen, zu Wahnvorstellungen und Größenwahn (Osler 2026), oder gar zu Gewalttaten und Suizid. Als Reaktion auf die Entdeckung, dass ein Attentäter in Kanada vor seiner Gewalttat diese mit dem Chatbot von OpenAI geplant hatte, hat sich OpenAI zwar entschuldigt[3] und kehrte zu einer früheren Version des Sprachmodells zurück, welches Nutzer:innen weniger um-

[2] https://www.wsj.com/tech/ai/mark-zuckerberg-ai-digital-future-0bb04de7
[3] https://www.deutschlandfunk.de/nach-toedlichen-schuessen-an-schule-openai-entschuldigt-sich-bei-kanadischer-gemeinde-fuer-nicht-gem-100.html

schmeichelt.[4] Allerdings löste dies Widerstand unter den Nutzer:innen aus – diese bevorzugten die schmeichlerische Variante. Der Spalt zwischen dem normativ Gebotenem und dem empirisch Bevorzugtem ist unübersehbar.

4.3. KI und Demokratie

Und wenn wir uns schon an die Reibungslosigkeit der Surrogate für Lernen und zwischenmenschliche Beziehungen gewöhnt haben, müssen wir uns dann wirklich noch mit der mühsamen Reibung der Demokratie beschweren? KI im Allgemeinen und Generative KI im Speziellen versprechen uns auch, die Mühsal der Demokratie zu überwinden. In einer Minimalversion verspricht man sich von Generativer KI, dass diese Dissense abbilden und abmildern kann, bspw. durch Moderation auf Online-Plattformen. Zudem könnte man aber über diese Moderationsfunktion hinaus versucht sein, die zwischenmenschlichen Dissense gleich aufzulösen, indem die KI auf Basis der Dissense entscheidet – als gleichsam höhere Vernunft, welche alle Argumente gewogen und basierend darauf neutral und objektiv entschieden hat. Delegation und Berechnung als Alternative zu Mühseligkeit und Schwerfälligkeit politischer und demokratischer Partizipation, Verhandlungen und Beratungen. Diese Verheißungen der Algokratie[5] haben ihre Vorläufer bereits in der Geschichte der Statistik (Porter 1995, Desrosières 1998) und den Versprechungen des evidenz-basierten Regierens, bekommen aber durch Generative KI noch mal Auftrieb: vom Auflösen der Dissense bis zur Auflösung des Menschen selbst in transhumanistischen Visionen.

Jenseits dieser Utopien oder Dystopien stellt insbesondere Generative KI bereits heute eine neue oder zumindest verschärfte Gefahr für individuelle und öffentliche Meinungsbildung und die demokratische Partizipation dar. Dies umfasst die Einfachheit und Schnelligkeit, mit der Desinformation oder Fake News hergestellt und geteilt werden können. Darüber hinaus können Personalisierungsmechanismen Sozialer Medien wie Generativer KI dazu führen, dass es insgesamt zu einer Reduktion von Reibung und Friktion mit Mitmenschen anderer Meinung kommt. Generative KI als personalisierter Filter zu jedweder Information könnte zu einer Verstärkung von Echokammern oder Filterblasen führen (Nguyen 2020), welche unsere Fähigkeiten, Dissense zu erleben, zu verstehen und mit ihnen umzugehen, unterminieren können.

Der desinformierte, aber verängstigte und isolierte Bürger ist ein ideales Opfer für die Verführung von Populismus und Faschismus. Dieser Verführung erliegen auch jene, welche an den Schaltstellen technologischer, politischer und ökonomischer Macht sitzen. Dies sollte einen sorgen, aber nicht wundern. Die Mächtigen brauchen keine demokratischen Verfahren, um die Vorteile demokratischer Gesellschaften genießen zu können. Sie haben Freiheit und Wohlstand, ohne Einschränkungen ihrer Macht durch demokratische Prozesse zu unterliegen. Es wird also darum gehen, wie sich die anderen 99% verhalten und ob sie

[4] https://openai.com/index/sycophancy-in-gpt-4o/

[5] Der Begriff Algokratie bezeichnet eine Form des Regierens, welche durch Computerprogramme strukturiert und organisiert ist. (vgl. Danaher 2016)

der Verführung der Reibungslosigkeit erliegen oder für Demokratie, Freiheit und Gleichheit kämpfen.

Literatur

Anderson, Chris. *What Science can learn from Google?* (Interdisciplines.org), 27. Dezember 2008.

Boyd, Danah, und Kate Crawford. „Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon“. *Information, Communication, & Society* 15, Nr. 5 (2012): 662–79.

Cheng, Myra, Cinoo Lee, Pranav Khadpe, Sunny Yu, Dyllan Han, und Dan Jurafsky. „Sycophantic AI decreases prosocial intentions and promotes dependence“. *Science* 391, Nr. 6792 (2026): eaec8352. https://doi.org/10.1126/science.aec8352.

Danaher, John. „The Threat of Algocracy: Reality, Resistance and Accommodation“. *Philosophy & Technology* 29, Nr. 3 (2016): 245–68. https://doi.org/10.1007/s13347-015-0211-1.

Desrosières, Alain. *The Politics of Large Numbers: A History of Statistical Reasoning*. Harvard University Press, 1998.

Hacking, Ian. „Statistical Language, Statistical Truth and Statistical Reason: The Self-Authentification of a Style of Scientific Reasoning.“ In *Social Dimensions of Science*, herausgegeben von Ernan McMullin. Notre Dame, Indiana, 1992.

Kitchin, Rob. „Big Data, New Epistemologies and Paradigm Shifts“. *Big Data & Society* April-June 2014 (2014).

Minsky, M. L. *Semantic Information Processing*. MIT Press, 1968.

Nguyen, C. Thi. „Echo Chambers and Epistemic Bubbles“. *Episteme* 17, Nr. 2 (2020): 141–61. https://doi.org/10.1017/epi.2018.32.

Osler, Lucy. „Hallucinating with AI: Distributed Delusions and “AI Psychosis”“. *Philosophy & Technology* 39, Nr. 1 (2026): 30. https://doi.org/10.1007/s13347-026-01034-3.

Porter, Theodore M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press, 1995.

Rieder, Gernot, und Judith Simon. „Datatrust: Or, The Political Quest for Numerical Evidence and the Epistemologies of Big Data“. *Big Data & Society*, Online-Vorab-Publikation, 2016. https://doi.org/10.1177/2053951716649398.

Turing, Alan. „Computing Machinery and Intelligence“. *Mind* 59, Nr. 236 (1950): 433–60.

Weiser, Mark. „The computer for the 21st century“. *SIGMOBILE Mob. Comput. Commun. Rev.* (New York, NY, USA) 3, Nr. 3 (1999): 3–11. https://doi.org/10.1145/329124.329126.

Von der Effizienz-Illusion und Bot-Delusion in Forschung und Lehre

Isabel Steinhardt, Christian Schneijderberg und Oliver Wieczorek[1]

Gemäß der Vision von Dario Amodei (2024), CEO von Anthropic, nimmt starke Künstliche Intelligenz eine fast messianische Gestalt an und wird als „Machines of Loving Grace" bezeichnet. Er geht davon aus, dass mit dem Aufkommen starker KI der Fortschritt in Biologie, Neurowissenschaften und Ökonomie innerhalb von fünf Jahren so groß sein wird wie der gesamte Fortschritt der Menschheit im 21. Jahrhundert. Darüber hinaus prophezeit er nicht nur technologische Durchbrüche, sondern auch die Realisierung eines ewigen Friedens sowie die Wiederherstellung eines neuen Lebenssinns, insbesondere in einer von KI gesteuerten Arbeitswelt.

Diese Vision bezieht sich nicht auf die derzeitigen Systeme der schwachen KI, sondern illustriert die Versprechen, die Tech-Firmen in die Welt setzen. Ein zentraler Baustein dabei ist stets das Effizienzversprechen, wenn KI nicht mehr nur als Hilfswerkzeug agiert, sondern als eigenständige Agenten, die beispielsweise ganze Forschungsprozesse übernehmen und menschliche Arbeit vollständig ersetzen. Das Versprechen ist dabei doppelt geschnürt: Einerseits suggeriert es eine Entlastung von repetitiven Aufgaben, andererseits eine Steigerung der kognitiven Kapazitäten. In der Forschung wird dies mit der Verheißung verbunden, dass KI Muster in riesigen Datensätzen erkennen kann, die dem menschlichen Auge entgehen, und dass der Forschungszyklus von der Hypothesenbildung bis zur Publikation drastisch verkürzt wird. In der Lehre wird das Versprechen etwa einer personalisierten Bildung und der Demokratisierung von Wissen durch adaptive Tutoren-Systeme in den Raum gestellt, die jedem Studierenden individuelle Betreuung auf Universitätsniveau bieten könnten.

Warum fällt dieses Narrativ auf fruchtbaren Boden? Die Antwort liegt in der tiefen strukturellen Krise des Wissenschaftssystems. Universitäten stehen seit Jahrzehnten unter massivem Performanzdruck in Bezug auf Veröffentlichungen („Publish or Perish") und die Akquise von Drittmitteln. Dazu kommen schlechte Betreuungsrelationen, knapper werdende Ressourcen, unsichere Arbeitsbedingungen im wissenschaftlichen Mittelbau sowie ein wachsender Vertrauensverlust der Gesellschaft in die Wissenschaft (Wissenschaftsrat, 2025, 2026). Dies führt bereits jetzt zu einer Krise der Wissenschaft, die sich in Phänomenen wie P-Hacking[2], unklaren wissenschaftlichen Praktiken, nicht-replizierbaren Studien und Fälschungen manifestiert (Fanelli, 2009; Ioannidis, 2014; Simmons et al., 2011).

[1] Die Gedanken zu Bot-Delusion beruhen auf umfassenden Diskussionen und einer gemeinsamen Publikation mit Christian Schneijderberg und Oliver Wieczorek, weshalb diese hier als Autoren mit aufgeführt sind.

[2] P-Hacking bezeichnet die manipulative Praxis in der Forschung, Daten oder statistische Auswertungen so lange anzupassen, bis ein nicht-signifikantes Ergebnis (p-Wert \(\ge \) 0,05) künstlich signifikant wird ($p < 0,05$).

In diesem Kontext erscheinen KI-Systeme als der scheinbar einfache Ausweg aus einem Dilemma, das eigentlich systemimmanent ist. Die Nutzung von KI verspricht mehr Output bei weniger Input. Geschieht diese Nutzung jedoch unkritisch, um die Produktivität zu steigern, ohne die ethischen und epistemologischen Implikationen zu hinterfragen, ist eine „Bot-Delusion" (Wieczorek et al., 2025) zu befürchten. Gemeint ist damit die unkritische Übergabe von Arbeitsprozessen in Forschung und Lehre an KI-Systeme bei unkontrollierter Übernahme der Outputs. Dabei entstehen nicht-intendierte Folgen (Merton, 1936), da KI-Systeme nicht neutral sind, sondern bestehende strukturelle Ungleichheiten verstärken oder abschwächen können, wie beispielsweise den Matthäus-Effekt (Merton, 1968).

Zudem ist problematisch, dass wir es bei KI-Systemen auf Basis von Large Language Models (LLMs) nicht mit Systemen zu tun haben, die versuchen, Zusammenhänge zu verstehen und kontextsensibel zu erklären. Vielmehr wurden diese Systeme darauf trainiert, Outputs zu erzeugen, die menschenähnlich klingen, und die auf Wahrscheinlichkeitsberechnungen beruhen. Deshalb wird von „Bullshit" gesprochen (Hicks et al., 2024), da LLMs keine wissenschaftlichen Erkenntnisse produzieren, sondern das Offensichtliche basierend auf den Trainingsdaten ausgeben (Dröge, 2025). Implizit bedeutet dies, dass die Trainingsdaten, auf denen die Berechnungen beruhen, zentral sind – wobei diese bei geschlossenen Modellen leider nicht bekannt sind. Wie sich diese Aspekte bei zukünftiger starker KI oder den derzeit aufkommenden Agenten zeigen werden, ist unklar.

Gleichzeitig werden aktuell kleinere LLMs entwickelt, die Open Source sind, deren Trainingsdaten bekannt sind und die lokal auf dem eigenen Rechner laufen können. Da diese Modelle in ihrer Leistungsfähigkeit zunehmen, könnten sie eine Alternative darstellen, die ethischen und datenschutzrechtlichen Anforderungen entspricht und damit für Forschung und Lehre besser nutzbar ist. Allerdings erfüllen diese Modelle nicht die großen Effizienzversprechen, sondern bleiben Arbeitswerkzeuge, da sie nicht leistungsstark genug sind, um Arbeitsprozesse vollständig zu ersetzen. Dies wird sehr wahrscheinlich dazu führen, dass sie zu einem Nischenphänomen verkommen. Im Folgenden wollen wir anhand von drei Beispielen aufzeigen, wo Effizienz-Illusionen und Bot-Delusions aufscheinen und welche Effekte diese haben könnten.

Literaturrecherche und Manuskripterstellung

Es gibt diverse Angebote wie Scite, Jenny AI oder Scopus AI für Wissenschaftler*innen sowie z. B. Hesse.AI für die Erstellung von Hausarbeiten. All diese Angebote versprechen Effizienzgewinne durch die Erstellung ganzer Manuskripte, die auf Literaturrecherchen mit echten Zitationen basieren. Damit unterscheiden sie sich von der „einfachen" Nutzung von LLMs wie ChatGPT, bei denen oft falsche Zitationen ausgegeben werden. Dabei unterschätzen Nutzer*innen die Natur sogenannter „Halluzinationen". Anders als ein menschlicher Fehler, der durch Nachlässigkeit entsteht, produzieren LLMs plausible, aber faktisch falsche Aussagen mit hoher sprachlicher Kohärenz. Diese „epistemische Unsicherheit" ist besonders tückisch: Die KI gibt nicht nur falsche Daten aus, sondern tut dies so überzeugend, dass sie für die menschliche Überprüfung kaum noch als fehlerhaft erkennbar ist.

Forschende und Studierende müssen bei der Nutzung der oben angegeben Angebote lediglich einen Prompt mit Forschungsfrage und Kontext eingeben (und bezahlen). Auch wenn nicht ganze Manuskripte erstellt werden, erscheint der Zeitgewinn durch die Literaturrecherche enorm. Hinzu kommen derzeit erste Angebote, bei denen die Literaturrecherche KI-Agenten überlassen wird, die – so das Versprechen – tausende Papers in Sekunden scannen und die relevantesten zusammenfassen.

Unklar bleibt bei all diesen Angeboten jedoch, wie die Auswahl der Literatur vorgenommen und nach welchen Kriterien gewichtet wird: nach Bekanntheit der Autor*innen, nach Zitationsindex oder nach Inhaltsübereinstimmung? Da dies unbekannt ist, können die Ausgaben nicht hinreichend bewertet werden, was zu nicht-intendierten Effekten führen kann. Auf diese Weise könnten neuartige Forschungsergebnisse, die zwar für die eigene Fragestellung hoch relevant sind, da sie vom Mainstream abweichen, aber schlecht zitiert wurden, ausgefiltert werden. KI-Systeme erkennen diese nicht als solche, da sie nicht kontextsensibel agieren, sondern nach algorithmischen Logiken arbeiten. Wenn die Literaturrecherche, die durch KI-Systeme erstellt wurde, unreflektiert übernommen wird und man der Illusion eines umfassenden Überblicks erliegt, kann es zum Verlust diverser Sichtweisen kommen. Dies ist der Fall, wenn KI-Systeme die Suche nach unkonventionellen Quellen als ineffizient markieren. Dadurch verliert Forschung ihre Innovationskraft, und Studierende verlernen wichtige Kompetenzen im Umgang mit Literatursuche und Schreibkompetenzen (Deskilling).

Experiment-Planung und -Durchführung

Dario Amodei beschreibt in seiner Vision starker KI, dass diese selbstständig Experimente plant und durchführt, ganz ohne menschliches Eingreifen. Bisher werden KI-Systeme lediglich für einzelne Aspekte wie die Hypothesenerstellung oder die Optimierung von Versuchsdesigns genutzt. Sehr wahrscheinlich greifen die KI-Systeme vor allem auf erfolgreiche Experimente zurück, da unerfolgreiche Experimente im Wissenschaftssystem kaum publizierbar sind (da keine signifikanten Ergebnisse vorliegen). Da KI-Systeme nicht selbstständig denken und nicht innovativ sind, werden experimentelle Ansätze, die erst erprobt werden müssen, aber häufig zu neuen Erkenntnissen führen, nicht vorgeschlagen. Dies führt zu einer Verstärkung des Status Quo.

Darüber hinaus gefährdet die algorithmische Optimierung den Verlust von Forschungsintuition und zufälligen Entdeckungen. Wissenschaftlicher Fortschritt entsteht oftmals durch das Verlassen bekannter Pfade und das (unbeabsichtigte) Abweichen von der Norm oder durch glückliche Zufälle. KI-Systeme, die darauf trainiert sind, Wahrscheinlichkeiten zu maximieren und Abweichungen zu minimieren, filtern genau diese als ineffizient heraus. Wer sich vollständig auf die KI verlässt, eliminiert systematisch die Chance auf Entdeckungen, die außerhalb der statistischen Wahrscheinlichkeit liegen. Dies kann langfristig (bei einer Bot-Delusion) zu einer wissenschaftlichen Stasis führen, bei der radikale Entdeckungen, die oft durch Fehler oder unerwartete Ergebnisse entstehen, systematisch unterdrückt werden. Diese Stasis wird verstärkt, wenn im Peer-Review KI-Systeme für die erste Sichtung eingesetzt werden und Mainstream-Artikel systematisch bevorzugt werden.

Lernen und Kompetenzentwicklung

Die Übertragung der „Bot-Delusion" auf den Bereich Studium und Lehre offenbart eine paradoxe Sackgasse. Zunächst würde die vollständige Auslagerung von Denk- und Schreibaufgaben zu einem systematischen Deskilling der Studierenden führen. Anstatt Kompetenzen durch den mühsamen, aber notwendigen Prozess des Übens und der Fehlerkorrektur aufzubauen, werden Studierende zu passiven Konsumenten von KI-Ergebnissen. Sie entwickeln eine trügerische Illusion der eigenen Kompetenz: Sie können Ergebnisse vorweisen, verstehen aber die dahinterliegenden Mechanismen nicht mehr: Das System produziert Absolvent*innen, die zwar effizient „produzieren", aber die Fähigkeit zur eigenständigen Problemlösung und kritischen Reflexion verloren haben.

Diese Bot-Delusion entsteht dann, wenn Lehrende ebenfalls Aspekte der Lehre an KI abgeben. Das heißt, sie bewerten die Produkte der Studierenden nicht kritisch, steigen nicht selbst in die kritische Betrachtung von KI-Systemen ein und verstehen daher nicht, wie KI-Erzeugnisse aufgebaut sind. Zudem stellen sie keine Aufgaben, die mit KI-Systemen (derzeit noch) nicht produzierbar sind, oder lassen Aufgaben und Produkte der Studierenden sogar durch KI-Systeme erstellen oder bewerten. Dadurch entsteht ein geschlossener Kreislauf der Selbstreferenz von KI-Systemen in der Aufgabenstellung, -bearbeitung und -bewertung. Es findet keine echte Prüfung von Wissen statt, sondern lediglich die Bestätigung statistisch wahrscheinlicher Antworten.

Ein weiteres zentrales ethisches und methodisches Problem ist die Undurchsichtigkeit dieser Systeme, oft als „Black Box" bezeichnet. Wenn KI-Systeme nicht nur Ergebnisse produzieren, sondern auch die Kriterien für deren Auswahl treffen, wird die Nachvollziehbarkeit wissenschaftlicher Prozesse unmöglich. In einer Welt, in der Entscheidungen über die Relevanz von Forschung oder die Bewertung von Studierenden durch undurchsichtige Algorithmen getroffen werden, bricht das Prinzip der wissenschaftlichen Rechenschaftspflicht zusammen. Wer haftet für ein falsches Urteil, wenn die Logik der Entscheidung nicht mehr in menschlicher Sprache erklärbar ist?

Parallel dazu könnte der flächendeckende Ersatz menschlicher Interaktion durch KI-Tutoren zu einer sozialen Entfremdung führen. Lernen ist ein sozialer Aushandlungsprozess, der auf Dialog, Widerspruch und der Begegnung mit dem Anderen beruht. KI-Systeme können zwar Feedback simulieren, basieren aber auf Wahrscheinlichkeiten und nicht auf echtem menschlichem Verständnis oder moralischer Autorität. Die Illusion der personalisierten Betreuung täuscht darüber hinweg, dass die essenzielle Komponente des Lernens eliminiert wird. Die Studierenden interagieren mit einem algorithmischen Spiegel, der keine echten Herausforderungen stellt, sondern nur das Bekannte reflektiert. Gleichzeitig entfremden sich Lehrende von den Studierenden und von wichtigem Input für die eigene Weiterentwicklung von Forschung und Lehre, wenn sie nur noch Wissen bereitstellen und kuratieren, aber nicht mehr aktiv beibringen sollen.

Um dieser Entwicklung entgegenzuwirken, bedarf es klarer Richtlinien, die den Einsatz von KI nicht als Selbstzweck zur Effizienzsteigerung, sondern als unterstützendes Werkzeug unter strenger menschlicher Aufsicht definieren. Dazu gehören die Pflicht zur

Offenlegung von KI-Nutzungen in Publikationen, die Entwicklung von Prüfverfahren, die „KI-Resistenz" fördern (also Aufgaben, die nur durch menschliche Urteilsfähigkeit lösbar sind), und die Stärkung von Open-Source-Modellen, deren Trainingsdaten und Logik einsehbar sind. Nur durch eine solche kritische KI-Literacy kann verhindert werden, dass die Wissenschaft sich selbst in einer Blase aus statistischen Wahrscheinlichkeiten und algorithmischen Verzerrungen einschließt.

Literatur

Amodei, D. (o. J.). Machines of Loving Grace. Darioamodei. Abgerufen 6. Mai 2026, von https://www.darioamodei.com/essay/machines-of-loving-grace

Dröge, K. (2025). Why AI has a “proving the obvious” problem, and what we can do about it | Computer-Assisted Qualitative Data Analysis. Computer-Assisted Qualitative Data Analysis. https://blogs.surrey.ac.uk/caqdas/2025/05/27/why-ai-has-a-proving-the-obvious-problem-and-what-we-can-do-about-it/

Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. PLOS ONE, 4(5), e5738. https://doi.org/10.1371/journal.pone.0005738

Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. Ethics and Information Technology, 26(2), 38. https://doi.org/10.1007/s10676-024-09775-5

Ioannidis, J. P. A. (2014). How to Make More Published Research True. PLOS Medicine, 11(10), e1001747. https://doi.org/10.1371/journal.pmed.1001747

Merton, R. K. (1936). The Unanticipated Consequences of Purposive Social Action. American Sociological Review, 1(6), 894. https://doi.org/10.2307/2084615

Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. Science, 159(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychological Science, 22(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Wieczorek, O., Steinhardt, I., Schmidt, R., Mauermeister, S., & Schneijderberg, C. (2025). The Bot Delusion. Large language models and anticipated consequences for academics’ publication and citation behavior. Futures, 166, 103537. https://doi.org/10.1016/j.futures.2024.103537

Wissenschaftsrat. (2025). Personalstrukturen im deutschen Wissenschaftssystem | Positionspapier [Application/pdf]. 91 pages. https://doi.org/10.57674/J6SF-H296

Wissenschaftsrat. (2026). Wissenschaft in Deutschland—Perspektiven bis 2040 (S. 133 pages) [Application/pdf]. Wissenschaftsrat. https://doi.org/10.57674/NPVJ-9804

Ambivalenzen des technologischen Fortschritts

Christian Leineweber

Wenn es um Fragen nach einer humanen Hochschulbildung in Zeiten von Künstlicher Intelligenz geht, dann scheint ein Fokus auf das Zusammenspiel von Technologie und Effizienz die Diskussion nahezu reflexartig in zwei Lager einzuteilen. Auf der einen Seite kann dafür plädiert werden, dass Technologien im Bildungssystem der Gesellschaft eingesetzt werden sollten, um Lern- und Bildungsprozesse sowie ihre Rahmenbedingungen effizienter, aber vor allem auch innovativer gestalten zu können. Auf der anderen Seite liegen kulturreflexive Befürchtungen darüber nahe, dass ein technologisches Effizienzstreben den Raum für individuelle Suchbewegungen, offene Leerstellen oder kreative Momente schwinden lässt und folglich herrschaftsbesetzte Lebensformen bedingt, die bereits Herbert Marcuse in den späten 1960er-Jahren unter dem Stichwort des *eindimensionalen Menschen* ausdifferenzierte. Ganz egal, ob der eindimensionale Mensch lernt oder verlernt, ob er effizient oder ineffizient ist, problematisch erscheint er für Marcuse zuvorderst aufgrund eines *technokratischen Bewusstseins*, das in einer instrumentellen Vernunft begründet liegt. Bei Marcuse (1967, S. 19) heißt es pointiert: „Im Medium der Technik verschmelzen Kultur, Politik und Wirtschaft zu einem allgegenwärtigen System, das alle Alternativen in sich aufnimmt oder abstößt. Produktivität und Wachstumspotenzial dieses Systems stabilisieren die Gesellschaft und halten den technischen Fortschritt im Rahmen von Herrschaft."

Mit dieser lediglich kursorischen Gegenüberstellung verdeutlicht sich in erster Linie eins: die Frage, wann Fortschritt durch Technologien in Tendenzen der Regression umschlägt, stößt mittlerweile auf ein durchaus breites geisteswissenschaftliches Fundament. Gleichsam scheinen kritische Analysen des Fortschritts nicht immer den Kern unserer wettbewerbsorientierten Gegenwart zu treffen. Der vorliegende Essay stellt sich der hier aufkeimenden Ambivalenz in bildungstheoretischer Absicht. In diesem Rahmen gilt es zunächst, eine Brücke zwischen dem Zusammenspiel von Technologie und Effizienz sowie dem pädagogischen Leitbegriff der Bildung zu schlagen (I). Daran anschließend soll dargelegt werden, dass epistemologische (II) und psychologische (III) Dimensionen im Zeitalter des Digitalen zunehmend mit ökonomischen Dimensionen verschmelzen. Insofern beide Formen der Verschmelzung pathologische Tendenzen hervorbringen, liegt es nahe, nach alternativen Möglichkeiten einer humanen Hochschulbildung der Zukunft zu fragen. Ausblicksartig (und gänzlich spekulativ) entwirft der Essay zum Abschluss eine Perspektive, die dafür sensibilisiert, dass die Möglichkeiten einer alternativen Hochschulbildung in der digitalen Transformation liegen könnten (IV). Jedoch scheint dies nur denkbar, wenn die digitale Transformation selbst unter Gesichtspunkten verhandelt und exploriert wird, die sich jenseits von Technologisierungs- und Effizienzbemühungen verorten.

(I) Der begriffliche Zusammenhang zwischen Technologie und Effizienz legitimiert sich im Wesentlichen mit einem unerschütterlichen Glauben unserer spätmodernen Gegenwartsgesellschaft am Fortschritt. Als Garant für diesen Fortschritt gilt im Allgemeinen eine stete Maximierung individueller und gesellschaftlicher Produktivität sowie die Akkumu-

lation von Kapital (z.B. Jaeggi 2023, S. 21ff.). Unter solchen Voraussetzungen stehen pädagogische Bemühungen immer schon in einem untrennbaren Verhältnis zu gesellschaftlichen Produktivitätslogiken. Denn institutionell gerahmte Lern- und Bildungsprozesse erweisen sich in erster Linie dann als gesellschaftlich anschlussfähig, wenn sie die ‚Entfesselung der Produktivkräfte' (Marx) ankurbeln. „Die pädagogische Idee der Bildung", so bringt beispielsweise der Bildungstheoretiker Werner Sesink (2014, S. 23f.) den hier adressierten Sachverhalt auf den Punkt, „wurde nur denkbar, weil sie sich auf die Entfesselung einer frei von der Bindung an vorhandene Mittel (den Stand der Produktivkräfte) entwickelbaren, demnach bildsamen Arbeitskraft beziehen konnte". Und Niklas Luhmann (1996, S. 19) lässt mit einem systemtheoretisch motivierten Blick von außen nüchtern ergänzen: „Ihre wohl wichtigsten Ziele sucht die Erziehung (was auch immer die Pädagogik davon halten mag) im Verhältnis zum Wirtschaftssystem." Innerhalb dieser Gemengelage profilieren sich medienpädagogische Bemühungen immer schon als Funktionen des Kapitalismus, indem sie technische Innovationen als Bedingungen der Möglichkeiten und Grenzen pädagogischen Handelns anerkennen. Wegweisend dafür scheint unter anderem die mediendidaktische Prämisse, dass die Ermöglichung von Lern- und Lehrangeboten mit digitalen Technologien interessanter, individueller, selbstwirksamer, optimierter und damit in letzter Konsequenz produktiver gestaltet werden soll (z.B. Kerres 2018).

(II) Dass es gerade Luhmanns nüchterne Beobachtung ist, die aus pädagogischer Sicht nicht einfach hingenommen werden sollte, lässt sich im Weiteren maßgeblich unter bildungstheoretischen Gesichtspunkten verdeutlichen. Als instruktiv erweist sich in diesem Zusammenhang ein Verständnis von Bildung als Transformation des subjektiven Welt- und Selbstverhältnisses, welches voraussetzt, dass Menschen in der lernenden Auseinandersetzung mit bislang unbekannten Inhalten, unbeantworteten Fragen oder ungelösten Problemen zu einer Erweiterung bis dato erlernter Wahrnehmungs- und Deutungsmuster gelangen (z.B. Marotzki 1990; Koller 2012). Bildungsprozesse basieren so gesehen in ihrem Kern auf einem Umgang mit dem *Unbestimmten* (z.B. Allert & Asmussen 2017, im Anschluss an Marotzki). Entlang dieser epistemologischen Setzung ist nun grundsätzlich bemerkenswert, dass die sozial- und kulturwissenschaftliche Literatur der Gegenwart die primäre Funktion des technischen Prozesses der Digitalisierung in einer „praktischen Herstellung von *Bestimmtheit*" sieht (Nassehi 2019, S. 30; Herv. C.L.). Angedeutet ist damit, dass eine wichtige Vision des Digitalen in einer datenbasierten und demnach logisch strukturierten Welterschließung zu liegen scheint (Bächle 2016, S. 115). Ein wesentliches Charakteristikum dieser logischen Welterschließung liegt darin, dass sie für den menschlichen Umgang mit Unbestimmtheit neue technologische Prinzipien anbietet und auf diese Weise in kulturelle Deutungsmuster, soziale Praktiken und letztlich auch pädagogische Handlungsformen eingreift. Dass etwaige Eingriffe allerdings als höchst ambivalent zu bewerten sind, kann insbesondere anhand algorithmischer und KI-gestützter Lernsysteme exemplifiziert werden. Wer sich beispielsweise auf digitalen Lernplattformen stark von guten Ergebnissen und maschinellen Empfehlungen leiten lässt, fühlt sein Handeln womöglich bestätigt, ist jedoch in gewisser Weise technisch geleitet und folglich fremdbestimmt. Nicht nur digitale Maschinen behaupten sich damit zunehmend als verkörperte Nachahmungen des Menschen (siehe bereits Weizenbaum 1978, S. 66), sondern der Mensch selbst scheint

sich unter digitalen Bedingungen in seinen Bestrebungen immer stärker maschinellen Logiken anzugleichen (Paoli 2025, S. 50). Dies korrespondiert schließlich damit, dass sich auch Lern- und Bildungsprozesse immer stärker an optimierte und ökonomisierte Abläufe annähern, in denen die epistemologische Kategorie des Unbestimmten systematisch reduziert (mitunter: negiert) wird, in denen also die Lösung gegenüber dem Problem dominiert. Daran gebunden sind nicht zuletzt Tendenzen des Wettbewerbs, der Outputorientierung und des Leistungsdrucks, die längst Eingang ins Bildungssystem gefunden haben.

(III) Der gesamtgesellschaftliche Kontext, in den die hier skizzierten Tendenzen einzuordnen sind, kann im Weiteren maßgeblich anhand der Subjektivierungsform der Individualisierung exploriert werden. Grundsätzlich gilt das Recht auf Individualisierung als große Errungenschaft unserer Gegenwartsgesellschaft, mit dem vor allem den Individuen westlicher Gesellschaften das Privileg eingeräumt wird, das eigene Lebensmuster selbstständig wählen zu können (Taylor 1995, S. 8). Digitale Medien und Applikationen folgen diesem Anspruch, indem sie höchst individualisierte Nutzungs- und Kommunikationsformen ermöglichen. Spürbar ist dies auch in pädagogischen Handlungsfeldern, wo die Digitalisierung beispielsweise mit der Anpassung von Lehr- und Lernmaterialien an die heterogenen Bedürfnisse von Lernenden, mit der Flexibilisierung von Lehr- und Lernformen oder mit der Personalisierung von Lehr- und Lernumgebungen korrespondiert (de Witt 2019, S. 814). Angesichts dessen ist jedoch ebenso bemerkenswert, dass der anhaltende technologische Fortschritt der Gesellschaft und die damit verbundenen, hoch flexibilisierten sowie entgrenzten Lebensweisen auf Seiten der Individuen verstärkt pathologische Tendenzen hervorbringen, die sich empirisch unter anderem in Formen der ‚Erschöpfung' (Ehrenberg 2015) oder ‚Überforderung' (Fuchs, Iwer & Micali 2018) niederschlagen. Dass diese Entwicklungen auch die Hochschulen betreffen, zeigt zum Beispiel der *Gesundheitsreport 2023* der Techniker Krankenkasse (2023, S. 16), in dem festgehalten wird, dass sich der Anteil von Studierenden, die sich gemäß ihrer Selbsteinschätzung gestresst fühlen, im Zeitraum von 2015 bis 2023 „von 23 Prozent auf 44 Prozent fast verdoppelt" hat. Insgesamt geht aus der Datenauswertung des Reports hervor, „dass an Hochschulen dringend mehr für die Gesundheit – vor allem für die psychische Gesundheit – getan werden muss" (ebd., S. 30).

An den damit offenkundig sichtbaren Pathologien der Gegenwart scheinen sich digitale Medien und Applikationen gemäß des Soziologen Steffen Mau (2018, S. 54) insofern direkt zu beteiligen, als insbesondere datenbasierte Bewertungen und individualisierte Rückmeldungen einen leistungssteigernden und motivationsfördernden Effekt auslösen; denn sie führen verstärkt dazu, dass „Menschen sich durch negative Vergleichsresultate in ihrem Selbstwertgefühl bedroht sehen und sich dann bemühen, beim nächsten Mal besser abzuschneiden". Dabei ist es gerade dieses als Bedrohung empfundene Gefühl, dass auch pädagogisch hoch bedeutsam scheint. So machen z. B. Forschungsergebnisse der empirischen Schulforschung darauf aufmerksam, dass die in deutschen Grundschulen zur individuellen Förderung der Lesekompetenz eingesetzte Lernplattform *Antolin* zu gesteigerten Formen des selbstoptimierenden Vergleichs unter Schüler:innen führt, selbst dann, wenn die Punktestände bewusst nicht im Unterricht thematisiert werden (Förschler et al. 2021,

S. 64). Angedeutet wird so, dass der Einsatz digitaler Medien in Lehr-/Lernszenarien die Verknüpfung von ökonomischen und psychologischen Dimensionen impliziert. In diesem Kontext wird die zu Beginn des vorliegenden Essays zitierte, kulturpessimistische Haltung Marcuses derzeit unter anderem eindrucksvoll durch den italienischen Philosophen Franco Berardi reaktualisiert. Berardi zufolge führt die gesellschaftlich ubiquitäre „Verwendung von Algorithmen" zu einer „logischen Reduktion", wobei es in erster Linie die Effizienz- und Fortschrittslogiken von Techniken sind, die subjektive Daseinsformen in eine „kapitalistische [...] Form" einlassen (Berardi 2019, S. 69). Entscheidend für Berardi ist dabei im Besonderen, dass sich die instrumentelle Herrschaft im Digitalen zunehmend auf psychische und kognitive Prozesse ausweitet, da Algorithmen weitestgehend im Verborgenen operieren: „Im Sozialen werden Automatismen ausformuliert, doch der Raum, in dem sie produziert werden, wird der Sichtbarkeit entrissen [...]. Nicht der Körper, sondern die Seele wird zum Subjekt technosozialer Herrschaft" (ebd., S. 231f.). Subjektive Wünsche, Bedürfnisse und Überzeugungen werden so von maschinellen Prozessen ‚geformt', die diesen vorausgehen.

(IV) Die diagnostizierten pathologischen Tendenzen unserer digitalen Gegenwartsgesellschaft mögen auf den ersten Blick alarmierend erscheinen; sie verlieren jedoch einen Teil ihrer Irritationskraft, sobald die Sozialgeschichte digitaler Medien systematisch in die Analyse einbezogen wird. Paradigmatisch dafür lässt sich beispielsweise die Monografie *Das Auge des Meisters* des italienischen Wissenschaftsphilosophen Matteo Pasquinelli (2024) anführen, welche die Entwicklung künstlicher Intelligenz als eine historische Fortschreibung von Arbeitsprozessen und Kontrolltechniken rekonstruiert. Wo die Entwicklung Künstlicher Intelligenzen folglich aus einem Bemühen um die Intensivierung von Produktivtät und Effizienz sowie der Steuerung und Kontrolle von Arbeitsprozessen hervorgeht, dort mag es schließlich deutlich weniger überraschen, wenn digitale Applikationen auf der Basis von Codes, Algorithmen und maschinellen Empfehlungen im Bildungswesen zu Kontroll- und Wettbewerbslogiken führen. Für eine humane Hochschulbildung in Zeiten von Künstlicher Intelligenz scheint es daher hilfreich, sensibel für Formen des medienpädagogischen Doings (Hofhues & Schütz 2022) zu werden, das heißt für konkrete Handlungen, Routinen und Praktiken, in denen sich digitale Medien als handlungs- und erfahrungsleitend erweisen.

Eine damit intendierte Verknüpfung von technischen Innovationen einerseits und pädagogischen Handlungen andererseits adressiert zunächst die bereits thematisierte mediendidaktische Zielvorstellung, dass sich mit technischen Innovationen die Gestaltung von Lehr- und Lernangeboten verbessern lässt. Diese Zielvorstellung setzt grundsätzlich aktive Lernende voraus, die ihr Wissen mithilfe medialer Angebote erweitern. Gerade in diesem Zusammenhang bleibt anzumerken, dass Formen der Aktivierung im Digitalen immer wieder neu gesetzt und transformiert werden. Sollte diese Einsicht einen richtigen Kern treffen, dann muss ihr die Ernüchterung auf dem Fuße folgen, dass wir den Wandel zu einer humanen Hochschulbildung in Zeiten von Künstlicher Intelligenz nicht mit tradierten Theorien, Annahmen und Konzepten gestalten können. Vielmehr gilt es empirisch sensibel für die neuen Praktiken, Wissensformen, Routinen und Irritationen zu werden, die digital

konstituiert werden, die also durch die Wirklogiken der digitalen Transformation selbst immer schon kontextualisiert und beeinflusst werden.

Für die Hochschulbildung erfordert dies nicht nur Vorstellungen davon, wie Student:innen lernen sollen; vielmehr scheinen Erforschungen und Reflexionen darüber notwendig, wie Student:innen (im Rahmen einer lebensweltlich etablierten Selbstverständlichkeit) lernen. Adressiert werden auf diese Weise solche Lernbewegungen, die sich auch deshalb massiv transformieren, weil digitale Medien zunehmend aktiver werden und sich folglich verstärkt in Lern- und Bildungsprozesse von Individuen einschreiben. Demnach lautet die zu priorisierende Frage nicht etwa, wie digitale Technologien zu mehr Effizienz führen, sondern wie bereits existierende studentische Medienpraktiken, die sich womöglich noch fernab konkreter Effizienzlogiken entfalten, reflexiv und verantwortungsvoll in Hochschulbildungsprozesse zu integrieren sind. Die Konsequenz wäre eine Hochschulbildung, die nicht lediglich für den „Maschinenraum" des Digitalen ausbildet (bzw. sich an diesen anzupassen versucht), sondern ihre epistemischen und praktischen Grundlagen aus eben diesem Maschinenraum heraus kritisch rekonstruiert und gestaltet. Nicht die digitalen Medien, sondern die Medienpraktiken von Studierenden – und die darin enthaltenen (auf Unbestimmtheit basierenden) Fragen, Suchbewegungen, Absichten, Wünsche, Ziele, Sehnsüchte usw. – müssen demzufolge ins Zentrum eines Nachdenkens um eine humane Hochschulbildung der Zukunft rücken. Ihr Leitbild konstruiert sich nicht mit den Fähigkeiten von Technologien, sondern verlagert den Fokus auf die Fähigkeiten von Menschen, die sich unter den derzeitigen technologischen Bedingungen längst vielfältig transformiert haben. Insofern auf diese Weise eine Transformation von innen angekurbelt wird, ist zuletzt die Einsicht wegweisend, dass das technologisierte Effizienzstreben unserer spätmodernen Gesellschaft zu einer relationalen Kopplung von Epistemologie, Ökonomie und Psychologie führt und sich auf Seiten der Subjekte zunehmend durch Pathologien bemerkbar macht. Sofern man diese Pathologien in ihren Symptomatiken ernst nimmt, liefern sie die Reflexionsräume dafür, wann individuelle Medienpraktiken an ihre Grenzen geraten. Die Hochschule der Zukunft hat daher zunehmend die originären Grenzen des Menschen gegenüber den Leistungsprinzipien des Technologischen zu berücksichtigen.

Literatur

Allert, Heidrun/Asmussen, Michael (2017). Bildung als produktive Verwicklung. In: Allert, Heidrun/Asmussen, Michael/Richter, Christoph (Hrsg.). *Digitalität und Selbst. Interdisziplinäre Perspektiven auf Subjektivierungs- und Bildungsprozesse*. transcript, S. 27-68.

Berardi, Franco ‚Bifo' (2019). *Die Seele bei der Arbeit. Von der Entfremdung zur Autonomie*. Matthes & Seitz.

Bächle, Thomas Christian (2016). *Digitales Wissen, Daten und Überwachung zur Einführung*. Junius.

Ehrenberg, Alain (2015). *Das erschöpfte Selbst. Depression und Gesellschaft in der Gegenwart*. Aus dem Französischen von Manuela Lenzen und Martin Klaus. Campus.

Fuchs, Thomas/Iwer, Lukas/Micali, Stefano (Hrsg.) (2018). *Das überforderte Subjekt. Zeitdiagnosen einer beschleunigten Gesellschaft*. Suhrkamp.

Förschler, Annina/Hartong, Sigrid/Kramer, Anouschka/Meister-Scheytt, Claudia/Junne, Jaromir (2021). Zur (ambivalenten) Wirkmächtigkeit datengetriebener Lernplattformen. Analyse des ›Antolin‹-Leseförderungsprogramms. In: *MedienPädagogik. Zeitschrift für Theorie und Praxis der Medienbildung*, 44 (Data Driven Schools), S. 52-72.

Hofhues, Sandra/Schütze, Kontanze (Hrsg.) (2022). *Doing Research – Wissenschaftspraktiken zwischen Positionierung und Suchanfrage.* transcript.

Jaeggi, Rahel (2023). *Fortschritt und Regression*. Suhrkamp.

Kerres, Michael (2018). *Mediendidaktik. Konzeption und Entwicklung digitaler Lernangebote*. De Gruyter.

Koller, Hans-Christoph (2012). *Einführung in die Theorie transformatorischer Bildungsprozesse*. Kohlhammer.

Luhmann, Niklas (1996). Das Erziehungssystem und die Systeme seiner Umwelt. In: Luhmann, Niklas/Schorr, Karl-Eberhard (Hrsg). *Zwischen System und Umwelt. Fragen an die Pädagogik*. Suhrkamp, S. 14-52.

Marcuse, Herbert (1967). *Der eindimensionale Mensch. Studien zur Ideologie der fortgeschrittenen Industriegesellschaft*. Luchterhand.

Marotzki, Winfried (1990). *Entwurf einer strukturalen Bildungstheorie – Biographietheoretische Auslegungen von Bildungsprozessen in hochkomplexen Gesellschaften*. DSV.

Mau, Steffen (2018). *Das metrische Wir. Über die Quantifizierung des Sozialen* (3. Aufl.). Berlin: Suhrkamp.

Nassehi. Armin (2019). *Muster. Theorie der digitalen Gesellschaft*. C. H. Beck.

Pasquinelli, Matteo (2024). *Das Auge des Meisters. Eine Sozialgeschichte Künstlicher Intelligenz*. Aus dem Englischen von Karina Hermes. Unrast.

Paoli, Guillaume (2025). *Etwas Besseres als der Optimismus*. Matthes & Seitz.

Sesink, Werner (2014). Eine kritische Bildungstheorie der Medien. In: Marotzki, Winfried/Meder, Norbert (Hrsg.). *Perspektiven der Medienbildung*. Springer VS, S. 11-44.

Taylor, Charley (1995). *Das Unbehagen an der Moderne*. Übersetzt von Joachim Schulte. Suhrkamp.

Techniker Krankenkasse (2023). *Gesundheitsreport. Wie geht's Deutschlands Studierenden?* URL: https://www.tk.de/resource/blob/2149886/e5bb2564c786aedb3979588fe64a8f39/2023-tk-gesundheitsreport-data.pdf.

Weizenbaum, Joseph (1978). *Die Macht der Computer und die Ohnmacht der Vernunft*. Suhrkamp.

de Witt, Claudia (2019). Digitalisierung in der Lehre. In: *Forschung & Lehre 9/2019*. Herausgegeben vom Deutschen Hochschulverband, S. 814-815.

Nachweise

Bei der Erstellung des vorliegenden Essays wurden Textpassagen und Überlegungen aus früheren Veröffentlichungen des Verfassers wiederaufgenommen. Die damit verbundenen sprachlichen und argumentativen Überschneidungen ergeben sich aus der erneuten Kontextualisierung bereits entwickelter Gedanken. Der vorliegende Beitrag verfolgt jedoch eine eigenständige Fragestellung: Er untersucht das Potenzial einer bildungstheoretischen Perspektive sowie daran gekoppelter autonomie- und gesellschaftsanalytischer Positionen für eine Reflexion der Universität der Zukunft.

Leineweber, Christian (2024). Widersprüche des Fortschritts. Perspektiven einer medienpädagogischen Kapitalismuskritik. In: Dander, Valentin/Grünberger, Nina/Niesyto, Horst/Pohlmann, Horst (Hrsg.). *Bildung und digitaler Kapitalismus*. München: kopaed, S. 129-142.

Leineweber, Christian (2026a). Das autonome Subjekt und die digitale Kultur. Pädagogische Reflexionen eines ambivalenten Verhältnisses. In: *Vierteljahrsschrift für Pädagogik* 102(2026), S. 231-247.

Leineweber, Christian (2026b). Studieren im Medium der Krise. In: Schaper, Sabrina/Lange, Jennifer/Hofhues, Sandra (Hrsg.). *(Fern-)Studierende – Lebenswelten und Übergänge*. Beltz Juventa, S. 207-214.

Leineweber, Christian (2026c). Digitale Medien der Unvernunft. Konturen eines Perspektivwechsels. In: Helbig, Christian/Hofhues, Sandra (Hrsg.). *Politiken der Digitalisierung und Medienpädagogik: Interdisziplinäre Deutungsangebote*. transcript, S. 45-61.

Der ambivalente Blick in den KI-Spiegel – mit Bezugnahme auf Christian Leineweber

Daniel Rütten

In seinem Essay für den vorliegenden Band zeigt Christian Leineweber die Ambivalenz in der gesellschaftstheoretischen Einordnung der technologischen Entwicklung auf, die sich auf das Verhältnis von Hochschulbildung und KI-Systemen wie folgt übersetzen lässt: Einerseits habe die Integration von KI-Systemen in die Hochschulbildung das Potenzial, Lern- und Bildungsprozesse effizienter, individueller und innovativer zu gestalten. Die Hoffnung für diese positive Entwicklung ergibt sich aus dem weit verbreiteten Fortschrittsglauben westlicher Gesellschaften. Andererseits lässt sich die Integration von KI-Systemen aus gesellschaftskritischer Sicht als Ausdehnung von hegemonialen Logiken instrumenteller Vernunft, Kontrolle und Wettbewerb verstehen. KI-Systeme fungieren demnach als Vermittlungsinstanzen von Herrschaft, weshalb die Wünsche und Bedürfnisse, die Studierende mit ihren KI-Praktiken verfolgen, unter Verdacht stehen, maschinell und herrschaftlich geformt zu sein. Leineweber plädiert jedoch dafür, über diese tradierte gesellschaftstheoretische Kritik hinauszugehen und die tatsächlichen KI-Praktiken Studierender empirisch zu analysieren und entsprechend ernst genommen werden. Der vorliegende Essay setzt an Leinewebers Forderung an, studentische KI-Praktiken und die sich darin ausdrückenden Wünsche, Bedürfnisse, Fragen und Unsicherheiten ins Zentrum der Überlegungen für eine humane Hochschulbildung zu stellen.

Ich werde Leinewebers Forderung entsprechend weiterdenken und konkretisieren. Zum *Weiterdenken* beschreibe ich KI-Technologien – in Anlehnung an Shannon Vallors Buch *AI Mirror* – metaphorisch als Spiegel. Ziel ist es, aus dieser Metaphorik eine kritische Analyse von KI-Praktiken abzuleiten, die über das empirische Beschreiben hinausweist. Zum *Konkretisieren* ziehe ich die Ergebnisse einer bundesweiten Studierendenbefragung zum Schreiben mit KI heran, in der sich zeigt: KI-Praktiken sind auch für viele Studierende keine alltägliche Selbstverständlichkeit, sondern ebenfalls geprägt durch Ambivalenzen oder Ablehnung. In Anbetracht dieser Ergebnisse sollte sich eine humane Hochschulbildung mit denjenigen Studierenden solidarisieren, die auf verschiedene Arten mit KI und ihrer eigenen KI-Nutzung zu kämpfen haben.

Weiterdenken: KI-Spiegel als Analyse-Instrument

Nach Vallor (2024) können KI-Systeme insofern als Spiegel verstanden werden, als sie uns kollektiv und individuell widerspiegeln, was wir in sie hineingeben und was wir bewusst und unbewusst in ihnen sehen wollen. Auf die Hochschulbildung bezogen fungieren KI-Systeme als Projektionsflächen für Wünsche und Bedürfnisse von Studierenden und ebenso von Lehrenden, Verwaltungsmitarbeiter*innen und der Hochschulleitung. Aufgrund dieser Möglichkeit, KI-Systeme als Projektionsflächen zu nutzen, geht von ihnen eine Verführungskraft aus. Indem die an Hochschulbildung beteiligten Personen in KI-Spiegel projizieren, was sie wollen oder brauchen, können projizierte Wunschbilder sie

dazu verführen, sich von KI-Spiegeln abhängig zu machen und die eigene Handlungsfähigkeit an diese abzutreten. So könnten wir laut Vallor von diesem Spiegelbild gefangen werden – wie Narziss, der von seinem eigenen Spiegelbild im Wasser gefesselt ist.

Diese Verführungskraft ist, so Vallor (2024), besonders stark, wenn wir uns nicht unserer Selbst gewahr sind – wenn beispielsweise Studierende nicht wissen, was sie können und wofür sie studieren oder wenn Lehrende daran zweifeln, welche Bedeutung universitäre Bildung noch hat.[1] Je größer die Unsicherheit über unser Selbstbild und unsere Fähigkeiten ist, desto stärker fühlen wir uns zu den Projektionen im KI-Spiegel hingezogen. Daraus ergibt sich die Gefahr, dass wir uns den zugrundeliegenden Problemen nicht mehr zuwenden, um sie individuell zu überwinden oder kollektiv bessere Bedingungen hierfür zu schaffen. Anstatt ausgehend von einer solchen Selbstverständigung die Art und Weise unserer Interaktionen mit KI-Systemen zu gestalten, könnten wir einem Technosolutionismus anheimfallen – der Vorstellung, dass jedes komplexe soziale, politische oder individuelle Problem eine technologische Lösung hat.

Das Potenzial von KI-Spiegeln sieht Vallor (2024) darin, die mit ihnen verfolgten Praktiken auf Wünsche, Bedürfnisse, Unsicherheiten und Kompensationen hin zu analysieren. Diese Analyse scheint genau die Forderung von Leineweber zu treffen, studentische Praktiken ins Zentrum bildungstheoretischer Überlegungen zu stellen.[2] In Anbetracht des Wechselspiels zwischen kollektiven und individuellen Projektionen sollten die Projektionen Studierender allerdings aus ihren gesellschaftlichen Funktionen und institutionellen Rahmenbedingungen heraus verstanden werden. Im Fokus der Analyse von KI-Praktiken Studierender nach Vallor könnten entsprechend folgende Fragen stehen: Was sehen Studierende im KI-Spiegel? Welche Wünsche, Bedürfnisse und Unsicherheiten werden in ihren KI-Praktiken verhandelt? Und was lässt sich daraus lernen über notwendige Veränderungen in der Hochschulbildung?

Konkretisieren: Projektionen, Ambivalenzen und Ablehnung

Ein erster empirischer Blick in den KI-Spiegel von Studierenden lässt sich aus den Daten der bundesweiten Studierendenbefragung zum akademischen Schreiben mit KI ableiten, die wir am Schreibzentrum der Goethe-Universität Frankfurt am Main durchgeführt haben (Hoffmann/Göbel/Rütten/Schmidt 2026). Die Befragung umfasst quantitative Daten zu verwendeten KI-Systemen, Nutzungsweisen, Gründen für die Nutzung und Einstellungen zu KI im Schreiben sowie offene Freitextantworten zu Veränderungen im Schreibprozess, im Umgang mit Wissen und in den Maßnahmen, Verantwortung für Texte zu übernehmen. In der Diskussion des Berichts synthetisieren wir die Antworten und identifizieren dabei drei Hauptfaktoren, welche ca. 90% der 4.048 befragten Studierenden dazu bringt, KI-Technologien in ihre Schreibprozesse zu integrieren: Zeit- und Notendruck, Schreib- und

[1] Der vorliegende Essayband ist zumindest zum Teil als Einladung dazu zu verstehen, diese Unsicherheit im Selbstverständnis zu bearbeiten.

[2] Im vorliegenden Essay liegt der Fokus dementsprechend auf den Projektionen von *Studierenden* auf KI-Spiegel. Konstantinidis (2025) betrachtet in seiner Anwendung der Spiegelmetapher auf Bildungssysteme stärker mögliche Projektionen von Lehrenden und Konsequenzen für Bildungsinstitutionen insgesamt.

Motivationsschwierigkeiten sowie ein ungedeckter Austauschbedarf (ebd. S. 62). Diese Hauptfaktoren dienen als Ausgangspunkt der oben motivierten Analyse von KI-Praktiken.

Der *Zeit- und Notendruck* im Hochschulsystem bringt Studierende dazu, den Effizienz- und Beschleunigungsversprechen von KI-Spiegeln zu verfallen. Sie sehen in KI die Lösung für die Wettbewerbslogiken, die Bildungsprozesse an Hochschulen maßgeblich prägen. In Anbetracht zunehmender Armut unter Studierenden und immer stärker werdender Notwendigkeit für Lohnarbeit können sich viele Studierende potenziell produktive Ineffizienzen, die mit der Konfrontation von Schwierigkeiten und Unsicherheiten einhergeht, schlicht nicht mehr leisten. In folgendem Freitext drücken sich die Notwendigkeiten in besonders deutlicher Sprache aus:

Der Schreibprozess wird durch klare Priorisierungen durch die KI beschleunigt. "menschliche Schwächen" (Konzentrationsmangel, Wortfindungsstörungen, Durchhaltevermögen, Selbstzweifel) werden effektiv beseitigt.

Anstatt „menschliche Schwächen“ als Chancen für Wachstum oder als Ausdruck zugrundeliegender Probleme zu verstehen, sind sie im ökonomisierten Bildungssystem Hindernisse, die es zu beseitigen gilt. Die Wirkmacht des Effizienzversprechens zeigt also nicht nur auf, wie weit die gesellschaftlichen und institutionellen Rahmenbedingungen von einer humanen Hochschulbildung entfernt sind, sondern auch, wie sich das eigene Selbstverständnis im Bildungsprozess unter diesen Rahmenbedingungen formiert.

Im derzeitigen Hochschulsystem ist es für Studierende außerdem schwer, mit *Schreibproblemen und fehlender Motivation* produktiv umzugehen. In diesem Zusammenhang äußern einige Studierende, dass ihnen weder das wissenschaftliche Arbeiten beigebracht noch die Bedeutung des Schreibens für universitäre Bildung plausibel gemacht wurde, wie beispielsweise in diesem Freitext deutlich wird:

[Ich] hatte […] große Angst vor dem akademischen Schreiben, weil ich das Gefühl hatte, dass ich mich nicht wissenschaftlich genug ausdrücke. Ich habe das Gefühl, dass wir wissenschaftliches Schreiben eigentlich nicht wirklich beigebracht bekommen und zu wenig Unterstützung diesbezüglich erhalten. Durch KI-Tools habe ich ein besseres Gespür dafür bekommen.

Die Defizite im Hochschulsystem lösen also Ängste bei Studierenden aus, die mit KI-Systemen kompensiert werden. Die Kompensation mit KI-Spiegeln bietet den Studierenden die Unterstützung, die sie im System nicht finden, verhindert aber gleichzeitig eine Überwindung dieser Ängste durch die anstrengende Aneignung von akademischen Schreibkompetenzen und die dafür nötige eigene Sinnsuche, die universitäre Bildungsprozesse ausmachen.

Der dritte Hauptfaktor für den Blick in den KI-Spiegel ist ein ungedeckter *Bedarf nach Austausch.* Im Spiegel ist jederzeit ein Gegenüber zu finden, auch wenn gerade keine Mitmenschen erreichbar sind. Diese Projektion einer immer verfügbaren Austauschpartner*in auf den KI-Spiegel hängt mit verschiedenen Problemen zusammen. Studierenden fehlen offensichtlich Gelegenheiten, einen bewertungsfreien Austausch über ihre Schreibprojekte

zu finden. Hilfsangebote in diesem Bereich sind unterfinanziert und ein Austausch mit Kommiliton*innen nicht für jede*n Studierende*n denkbar oder möglich. Lehrende haben oft nicht die Kapazitäten, so zu beraten wie sich dies einige Studierende – ebenso viele Lehrende – wünschen würden. Dies wird sich durch die bereits vielerorts spürbaren Kürzungen in der Bildung vermutlich eher verschärfen. KI-Spiegel werden außerdem bewusst so gestaltet und beworben, dass Studierende sie als parasoziale Begleitung wahrnehmen und in ihnen die Bestätigung finden, die sie suchen (z.B. Sommerer 2026).

Im folgenden Freitext werden alle drei Hauptfaktoren in ihrem komplexen Zusammenhang beschrieben, weshalb er an dieser Stelle als Zusammenfassung dienen kann:

Ein "Gesprächspartner" im Sinne von KI ist sofort und immer zur Verfügung. Schreibprozess kann somit schneller gelingen, da in einem "Austausch" mit der KI die Kreativität angeregt werden kann und die Möglichkeit gegeben wird schneller auf neue Gedanken zu kommen. (Keine Notwendigkeit akut und bei Bedarf gezielte Personen aufzusuchen, die sich mit dem Bereich auskennen oder offen sind für wissenschaftliche Diskurse, um das eigene Denken anzuregen. Keine Notwendigkeit ein paar Stunden etwas anderes zu tun, um möglicherweise "den Knoten im Denken" während des Schreibprozesses zu lösen.).

Die Beweggründe für die KI-Nutzung weisen darauf hin, dass die Effizienzlogik erheblichen Einfluss auf Bildungsprozesse hat, Studierende zu wenig inhaltlich begleitet werden sowie sich insgesamt oft wünschen, mehr Austauschmöglichkeiten zu haben. All diese Probleme werden mit KI-Spiegeln bewusst und unbewusst kompensiert, was insgesamt wohl nur wenige überraschen wird. Kaum Beachtung findet die Tatsache, dass Studierende ebenso *ambivalente Haltungen* gegenüber dieser Entwicklung äußern wie die von Leineweber angeführten Gesellschaftstheoretiker*innen. Dies steht der weitverbreiteten Erzählung entgegen, dass Studierende KI-gestützte Praktiken als ausschließlich effizient, vorteilshaft und selbstverständlich empfinden.

Zieht man die hier eingeführte Spiegelmetapher heran, wird die Ambivalenz Studierender verständlich: Sie werden sich - stellenweise - der Verführungskraft von KI-Spiegeln bewusst. Studierende bemerken, dass sie die unzureichende Kompensation langfristig wieder einholt und sie damit ihre potenziell zu entwickelnden Fähigkeiten auf KI übertragen. Durch die extensive Nutzung fühlen sie sich letztlich nicht mehr in der Lage, ihre eigenen Kräfte noch zu nutzen. Sie beschreiben sich oftmals selbst als faul oder – wie in folgendem Freitext – abhängig von der KI:

Ich bemühe mich nicht so sehr beim Schreiben von wissenschaftlichen Texten und neige zu prokrastinieren dadurch. Ich bin auch weniger kreativ und habe weniger Ideen was und wie ich machen kann ohne KI. Ich muss sagen ich bin schon abhängig von ihr geworden.

Diesen Studierenden fällt es immer schwerer, sich von der Projektionsfläche zu lösen, die ihnen immer die einfache Lösung widerspiegelt. Sie merken, dass sie sich immer schlechter an Geschriebenes erinnern können oder sich von ihren Textprodukten entfremden. Sie versuchen zwar selbst, diesen Problemen entgegenzuwirken, indem sie beispielsweise ihre KI-Nutzung in bestimmten Phasen des Schreibprozesses einschränken. Allerdings fühlen sie sich mitunter ohnmächtig, der Verführung zu widerstehen, „wenn ChatGPT nur zwei

Klicks entfernt ist“[3]: Sobald einmal die KI genutzt wird ist es schwerer im weiteren Verlauf, nicht auf die KI zuzugreifen.

Die vielfältige, aber deutliche Schilderung von Ambivalenzen gegenüber der eigenen KI-Nutzung sollte Berücksichtigung finden bei der geforderten verantwortungsvollen Integration neuer KI-Praktiken in die Hochschullehre und in den entsprechenden Konzepten von *AI Literacy*. Studierenden sollten Möglichkeiten aufgezeigt werden, die Verführung von KI-Spiegeln erkennen und benennen zu können. Hilfreich könnten dabei Begriffe aus dem Forschungsfeld der Human-Computer-Interaction sein, welches sich schon länger mit dem Einfluss von automatisierten Systemen auf menschliches Handeln auseinandersetzt. Mithilfe von Begriffen wie Nudging, Choice Architecture, Gebrauchssuggestionen, Framing etc. lassen sich beispielsweise Analysen und Reflexionen von KI-Interaktionen der Studierenden anleiten. Aufbauend auf einem so erreichten Bewusstsein für die Verführung, sollte Studierenden Handlungsstrategien aufgezeigt werden, wie sie mit Verführungen praktisch umgehen. Und nicht zuletzt sollte auf Alternativen zur Kompensation hingewiesen werden, die es dafür teilweise aufzubauen, zu bewahren und zu stärken gilt. So könnten Studierende produktiv mit ihrer Ambivalenz umgehen und sich womöglich von der Verführung lösen.

Neben dieser unterbeleuchteten Gruppe an Studierenden, die ambivalent zu ihrer eigenen KI-Nutzung stehen, wird eine weitere Gruppe vielleicht sogar noch weniger beachtet: die 10% der Studierenden, die ohne KI-Spiegel schreiben. Die Freitexte aus der Studierendenbefragung zeigen auf, dass dieser Verzicht mittlerweile eine informierte Entscheidung ist, insbesondere weil diese Studierenden befürchten, ihre eigenen Fähigkeiten zu stark auf KI-Spiegel zu übertragen, aber auch, weil sie aus anderen ethischen Gründen eine KI-Nutzung ablehnen, z.B. aufgrund der Umweltfolgen, der Abhängigkeit von Tech-Konzernen oder der inhärenten Ausbeutung, ohne die KI-Systeme nicht möglich wären. Studierende betrachten ihre eigene KI-Verweigerung damit auch als eine Form des Widerstands. Auch diese Studierende sollten Beachtung finden in der Neuausrichtung der Hochschulbildung. Die Nutzung von KI-Systemen im Studium sollte eine Möglichkeit, keine Voraussetzung sein.

Ein kurzes Fazit

Der in diesem Essay skizzierte Blick in den KI-Spiegel von Studierenden zeigt, dass die Haltungen Studierender gegenüber ihren eigenen KI-Praktiken vielfältig sind. Eine verantwortungsvolle Integration von KI-Praktiken in die Hochschulbildung muss der Vielfalt dieser Haltungen Studierender gerecht werden. Diese Praktiken ins Zentrum der Überlegungen für eine humane Hochschulbildung zu stellen, bedeutet, die Beweggründe für die Praktiken zu verstehen, Ambivalenzen zu berücksichtigen und Ablehnung zuzulassen. Schon ohne eine tiefe gesellschaftstheoretische Interpretation zeigen sich in KI-Praktiken Defizite des Bildungssystems, die sich nicht allein auf technischer Ebene lösen lassen. Ein neues Defizit ergibt sich aus der Ohnmacht, die Studierende gegenüber der Verführungskraft von KI-Spiegeln wahrnehmen. Ablehnung gegen diese Praktiken ist gut begründet,

[3] Siehe zu dieser Verführung auch den Essay von Judith Simon im vorliegenden Band

aber wird nicht immer respektiert. Es erscheint für alle Beteiligten gleichermaßen wichtig, sich der Unsicherheit darüber zu stellen, was Hochschulbildung in Zeiten von KI ausmacht. Erst diese Unsicherheit macht die Verführungskraft von KI-Spiegeln so stark. Antworten dürfen allerdings nicht im hochschuldidaktischen Diskurs verbleiben, sondern sollten gemeinsam mit Studierenden diskutiert, erprobt und weiterentwickelt werden.

Literatur

Hoffmann, N.; Göbel, H.; Rütten, D.; Schmidt, S. (2026). „Manchmal muss ich mich aktiv zum selbst denken motivieren". Bericht zur Studierendenbefragung 2025 zum akademischen Schreiben mit KI. Schreibzentrum der Goethe-Universität Frankfurt a. M. https://doi.org/10.21248/gups.97289.

Konstantinidis, A. (2025). A Metaphor for Rethinking Artificial Intelligence in/and Education. Journal of Interactive Media in Education, 2025(1). https://doi.org/10.5334/jime.981.

Sommerer, T. (2026). Artificial Resonance: AI companions as agents of social acceleration. *AI & SOCIETY*. https://doi.org/10.1007/s00146-026-03179-3

Vallor, S. (2024). The AI mirror: How to reclaim our humanity in an age of machine thinking. New York, NY: Oxford University Press.

Veränderte Mensch-Technologie Relationen – Perspektiven auf multistabile und imitierende Technologien

Silke Elisabeth Wrede

Um einen Beitrag zu einer Idee der Universität des 21. Jahrhunderts zu leisten, werden im Folgenden Veränderungen in den Mensch-Welt Relationen skizziert, die insbesondere durch die zunehmende Verbreitung und Weiterentwicklung digitaler Technologien beeinflusst werden. Diese Veränderungen sind von Bedeutung, da sie die Zuschreibungen studentischer Leistungen, die Art und Weise des Lernens sowie deren Bewertung beeinflussen.

Im Anschluss an Don Ihde (1990) lassen sich Mensch-Welt Relationen als technologisch vermittelte Relationen begreifen und in hermeneutische (hermeneutic), verkörperte (embodiment), Alteritäts- (alterity) und Hintergrundrelationen (background) ausdifferenzieren. Jede neue oder weiterentwickelte Technologie kann neue Möglichkeiten eröffnen, die Welt wahrzunehmen, zu verstehen und zu erfahren. Hochschulen also Orte der Vermittlung von Welterfahrung (Schwan, 2015) sind davon unmittelbar betroffen, da aktuelle technologische Entwicklungen wie generative KI Prozesse des Deutens, Prüfens und der Wissensgenerierung beeinflussen können.

Wie sich die Veränderungen konkret zeigen, ist bislang offen, daher wird explorativ der Frage nachgegangen, wie sich die Relationen zwischen Mensch und Technologie im Zuge der Verbreitung generativer KI verändern und welche Auswirkungen dies für Lehren und Lernen an Hochschulen haben kann. Im Mittelpunkt stehen dabei zwei Phänomene die Multistabilität generativer KI-Tools und ihre Imitation menschlicher Fähigkeiten. Die Hintergrundrelation wird nicht näher betrachtet, da die aktive Nutzung im Fokus stehen soll.

Hermeneutische Relation

Die hermeneutische Relation beruht auf der Grundannahme, dass Technologien einen Weltausschnitt zeigen, den es erst zu lesen und zu deuten gilt (Ihde, 1990, S. 80 ff.). Das betrifft sowohl das Artefakt selbst als auch seine gesellschaftliche Bedeutung. Auch generative KI vermittelt Welt auf eine bestimmte Weise, da sie Informationen auswählt, ordnet und sprachlich darstellt. Daher sind es bei generativer KI vor allem die Ausgaben, die gelesen und gedeutet werden müssen. Sie angemessen zu deuten, setzt Wissen darüber voraus, wie diese Ausgaben zustande kommen, insbesondere über statistische Wahrscheinlichkeiten und Mustervergleiche. So kann generative KI beispielsweise eine sprachlich überzeugende Erklärung zu einem fachlichen Sachverhalt generieren, die bei genauerer Prüfung jedoch Fehler oder Vereinfachungen enthält. Die Nutzenden müssen somit die Ausgabe nicht nur verstehen, sondern auch auf ihre fachliche Richtigkeit hin beurteilen. Ergänzend sind Kenntnisse über Nutzungsweisen und das Prompting sowie mögliche Einsatzszenarien der Tools in der Lehre und deren Bedeutung für Lehr- und Lernprozesse wesentlich. Wie generative KI eingesetzt wird, beeinflusst nämlich auch, wie Kontrolle und Verantwortung für Lehr- und Lernaufgaben verteilt sind. So können Lehrende voll-

ständig für die Generierung und Überprüfung von Aufgaben zuständig sein, während die KI lediglich unterstützende Informationen bereitstellt. Alternativ können Lehrende den Einsatz der KI steuern und überwachen, während Studierende mit KI-generierten Aufgaben arbeiten (Molenaar, 2022, S. 634 ff.).

Gerade beim Einsatz in der Lehre zeigt sich eine Vielzahl von Nutzungsweisen der generativen KI, die sich mit dem Begriff der Multistabilität erfassen lassen. Multistabil meint, dass „any technology can be put to multiple purposes and can be meaningful in different ways to different users" (Rosenberger & Verbeek, 2015, S. 25). Einzelne generative KI-Tools können für verschiedene Lehr- und Lernzwecke eingesetzt werden und sind für verschiedene Nutzende auf unterschiedliche Art und Weise bedeutsam. So reicht die Spannbreite der Nutzungsweisen für die Textbearbeitung von Übersetzung und Textgenerierung über Brainstorming bis hin zu inhaltlichen Erklärungen. Dabei stellt sich für die Hochschulen die Frage, wie eine Technologie mehrere unterschiedliche Nutzungs- und Bedeutungsweisen annehmen kann und zugleich innerhalb bestimmter Anwendungsszenarien hinreichend stabil bleibt (Breil, 2025, S. 184). Eine solche Stabilität kann wichtig sein, damit Lernaktivitäten eingeübt, wiederholt und in ihrer Wirkung beurteilt werden können.

Neben der Deutung ihrer Ausgaben ist auch die Art der Interaktion mit generativer KI für Lehr- und Lernprozesse relevant. Durch die dialogische Gestaltung kann KI als kommunizierendes Gegenüber wahrgenommen werden. Damit rückt die Alteritätsrelation in den Blick.

Alteritätsrelation

In der Alteritätsrelation nach Ihde (1990, S. 97 ff.) wendet sich der Mensch der Technologie zu und interagiert mit ihr. Dabei wird die generative KI als Quasi-Anderer wahrgenommen, als von der eigenen Person verschieden, aber verbunden zum Beispiel in Frage-Antwort-Strukturen oder als schriftliche:r Kommunikationspartner:in. Laaksoharju et al. (2023) zufolge verändert sich insbesondere die Kommunikation mit den Technologien. Damit ergeben sich auch die Nutzungsweisen, denn generative KI ermöglicht dialogische Interaktionen, in denen Nutzende Rückfragen stellen und Antworten präzisieren. Dadurch greifen erlernte Muster und Strukturen zwischen Mensch und Maschine nicht mehr. Sprachkompetenz und kausale Argumentation galten lange als charakteristische menschliche Fähigkeiten, anhand derer Kompetenz, Vertrauen und Verlässlichkeit eingeschätzt wurden. Wenn Technologien diese Fähigkeiten überzeugend imitieren, können entsprechende Zuschreibungen durch die Maskierung der Andersartigkeit (siehe Antje Michels in diesem Band) auf die Technologie übertragen werden. Die technologische Andersartigkeit tritt hinter die Kommunikationsweise zurück, die menschlichen Ausdrucks- und Erklärungsmustern ähnelt. Generative KI-Technologien sind von Menschen so entwickelt worden, dass sie zufriedenstellende Antworten erzeugen. Sie tendieren dazu, Nutzenden zuzustimmen oder deren Erwartungen entgegenzukommen. Dabei orientieren sie sich an den individuellen Nutzenden und lassen deren Meinungen und Formulierungen mitunter als allgemeingültig erscheinen (in Anlehnung an Fan et al., 2026, S. 19 ff.). Die Kommunikation gestaltet sich somit reibungsloser als mit anderen Menschen, in der unterschiedliche

Sichtweisen aufeinandertreffen (siehe Judith Simons in diesem Band). Damit wird deutlich, dass die Gestaltung und Funktionsweise generativer KI bestimmte Formen der Kommunikation begünstigen und beeinflussen können, wie Nutzende die Antworten wahrnehmen und bewerten.

Das Infragestellen von vertrauten Mustern lässt sich an zwei weiteren Aspekten erkennen. Erstens können die Ausgaben von generativer KI gut strukturiert, überzeugend und intern konsistent gestaltet sein, so dass sie wie eine plausible Erklärung wirken. Allerdings ist dies bei genauerer Analyse mitunter nicht der Fall. Zweitens können Schritt-für-Schritt-Erklärungen der generativen KI den Eindruck vermitteln, der Weg zum Ergebnis sei transparent nachvollziehbar. Da solche Erklärungen menschlichen Erklärungsmustern entsprechen, können sie das Vertrauen in die Ausgabe verstärken (in Anlehnung an Fan et al., 2026, S. 19 ff.). Diese Imitationen können Fehlzuschreibungen begünstigen, die sich auch als Irritationen in der Mensch-Technologie Relation beschreiben lassen. Ihre Offenlegung ist zentral, damit die hermeneutische Relation gelingen kann. Zu klären ist, welchen Ausschnitt der Welt ein generatives KI-Tool vermittelt, über welche Fähigkeiten es tatsächlich verfügt, wie seine Gestaltung und Programmierung die Reaktionen auf Menschen prägen und welche Ziele die Entwicklung und Nutzung des Systems leiten.

Verkörperte Relation

Bei der verkörperten Relation wird die Technologie derart in eine Tätigkeit eingebunden, dass die Aufmerksamkeit nicht mehr primär auf das Artefakt, sondern auf die Tätigkeit und deren Ziel gerichtet ist. Ihde bezeichnet dieses in den Hintergrund treten als Transparenz, das Artefakt wird in der Tätigkeit transparent (Ihde, 1990, S. 72 ff.). Ein Beispiel ist die Tastatur, mit der eine Literaturrecherche durchgeführt wird. Zunächst muss der Umgang mit der Tastatur bewusst erlernt werden. Anschließend kann die Literaturrecherche in den Vordergrund treten, während die Tastatur in den Hintergrund tritt. Erst bei einer Störung, etwa durch eine falsch gedrückte Taste, wird sie wieder zum Gegenstand bewusster Aufmerksamkeit.

Was geschieht jedoch, wenn Technologien enger mit dem menschlichen Körper verschmelzen wie beispielsweise mit Implantaten oder XR-Brillen? Verbeek (2008; 2015) erweitert den Ansatz von Ihde um verschiedene Formen der Intentionalität. In der phänomenologischen Tradition bezeichnet Intentionalität ursprünglich die Gerichtetheit menschlicher Erfahrung auf die Welt. Ihde ergänzt dieses Verständnis von Intentionalität, indem er auch die Gerichtetheit technologischer Artefakte in den Blick nimmt, während Verbeek untersucht, wie Technologien menschliche Wahrnehmungs- und Handlungsbezüge zur Welt mitprägen.

In der verkörperten Relation richtet sich die menschliche Intentionalität durch ein technologisches Artefakt auf die Welt. Dabei kann es vorkommen, dass technologische Gestaltung die Wahrnehmung und Deutung der Nutzenden beeinflusst, ohne in der Interaktion unmittelbar sichtbar zu werden. Bei der hybriden Intentionalität wird die Technologie hingegen funktional in den Menschen integriert wie bei einem Implantat. Verbeek (2008, S. 392) beschreibt damit eine Verbindung menschlicher und technologischer Anteile, ohne

dass diese deshalb grundsätzlich ununterscheidbar wären. Komposite Intentionalität bezeichnet das Zusammenwirken menschlicher und technologischer Intentionalität. Wenn Studierende sich mit Hilfe generativer KI einen komplexen wissenschaftlichen Begriff erklären lassen wollen, dann verarbeitet das KI-System die Anfrage und stellt die Erklärung auf eine bestimmte Weise dar. Es wählt bestimmte Aspekte aus, gewichtet diese und generiert eine Antwort. Studierende lesen die Erklärung und entwickeln auf dieser Grundlage ein Verständnis des Fachbegriffs. Damit richtet sich die Person nicht auf den Fachgegenstand, sondern auf die Weise, in der KI ihn darstellt.

Die verschiedenen Formen der Intentionalität verdeutlichen, wie eng menschliche und technologische Anteile an Wahrnehmung und Handeln miteinander verbunden sein können. Auch die verschiedenen Mensch-Technologie Relationen lassen sich in konkreten Anwendungssituationen nicht immer eindeutig voneinander abgrenzen. XR-Brillen können zum Beispiel das unmittelbare Erleben der physischen Umwelt mit einer digitalen Darstellung verbinden. Während die Wahrnehmung der Umgebung der verkörperten Relation zugeordnet werden kann, müssen die eingeblendeten digitalen Informationen zunächst interpretiert werden. Hieran wird ersichtlich, dass die verschiedenen Relationen analytisch unterschieden werden können, in konkreten Mensch-Technologie Beziehungen jedoch immer mehr ineinandergreifen können.

Auswirkungen auf die Bildung

Neben den irreführenden Eindrücken, die aus der Imitation menschlicher Fähigkeiten in der Mensch-Technologie Relation entstehen, ermöglicht generative KI auch Möglichkeiten, die eigene Lernleistung hinter einem mithilfe der Technologie erstellten Arbeitsprodukt zu verbergen. Die Qualität eines eingereichten Textes lässt nicht mehr zwangsläufig darauf schließen, ob Studierende entsprechende Kenntnisse und Fähigkeiten erworben haben (Yan & Gašević, 2026). Studierende können somit gute Qualifikationsarbeiten einreichen, ohne im entsprechenden Umfang gelernt zu haben, da ein generatives KI-Tool Teile dieser Aufgabe übernommen hat. Die einfache Verfügbarkeit der Tools kann dies begünstigen und kurzfristig individuell lohnend sein. Allerdings stellt diese Entwicklung bisherige Lehr- und Lernpraxis in Frage. Wenn an einem eingereichten Arbeitsprodukt die Kompetenzen der Studierenden nicht mehr beurteilt werden können, dann verlieren produktorientierte Prüfungsformate an Aussagekraft.

Welche Konsequenzen diese Verschiebungen und Neuausrichtungen für den Bildungsbereich haben, ist noch offen. Die Analyse macht jedoch drei Spannungsfelder sichtbar. Erstens besteht ein Spannungsfeld zwischen der Vielfalt an möglichen Nutzungsweisen generativer KI und dem Bedarf an stabilen Lernwegen (Multistabilität). Zweitens stehen überzeugend wirkende Ausgaben und ihre tatsächliche Verlässlichkeit in einem Spannungsverhältnis (hermeneutische und Alteritätsrelation) und drittens besteht ein Spannungsfeld zwischen der Qualität des Arbeitsprodukts und dem individuellen Lernfortschritt.

Daraus lassen sich erste Ansatzpunkte für Hochschulen, Forschende, Lehrende und Studierende ableiten. Erstens gilt es, Multistabilitäten zu identifizieren, Räume zu schaffen, in denen diese praktisch erprobt und erfahren werden können, und zugleich verlässliche

Lernwege aufzuzeigen. Zweitens sollten vor der Interaktion mit den Technologien sowohl die eigenen Ziele als auch die Gestaltungsziele und systemseitigen Ausrichtungen der Technologie reflektiert werden. Drittens erfordert die Deutung der Ausgaben der generativen KI ein kritisch-analysierendes Vorgehen, das die Imitation menschlicher Fähigkeiten in den Blick nimmt. Viertens sollten Lehrende und Studierende den langfristigen Wert des Lernens gegenüber dem kurzfristigen Erfolg einer Qualifikationsarbeit stärker gewichten. Das schließt auch die Weiterentwicklung von Prüfungsformaten ein. Insgesamt sind mit den kommenden Entwicklungen wie der agentischen KI, die deutlich eigenständiger agiert, noch komplexere Mensch-Technologie Relationen zu erwarten, die neue Analysen nach sich ziehen werden.

Literatur

Breil, P. (2025). Ambiguous Relations: A Postphenomenological Reflection on Technological Multistability in Education. In M. Bohlmann & P. Breil (Hrsg.), *Postphenomenology and Technologies within Educational Settings* (S. 175–193). Lexington.

Fan, Y., Yan, L., Molenaar, I., Greiff, S. & Gašević, D. (2026). *A Metacognitive Approach to Learning and Performance in Human-AI Interaction.* Preprint. https://doi.org/10.31234/osf.io/s7dvk_v1

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. *The Indiana series in the philosophy of technology*. Indiana University Press.

Laaksoharju, M., Lennerfors, T. T., Persson, A. & Oestreicher, L. (2023). What is the problem to which AI chatbots are the solution? AI ethics through Don Ihde's embodiment, hermeneutic, alterity, and background relationships. In T. T. Lennerfors & K. Murata (Hrsg.), *Ethics and Sustainability in Digital Cultures* (S. 31–48). Routledge. https://doi.org/10.4324/9781003367451-4

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632–645. https://doi.org/10.1111/ejed.12527

Rosenberger, R. & Verbeek, P.-P. (2015). A Field Guide to Postphenomenology. In R. Rosenberger & P.-P. Verbeek (Hrsg.), *Postphenomenology and the Philosophy of Technology. Postphenomenological Investigations: Essays on Human-Technology Relations* (S. 9–41). Lexington Books.

Schwan, G. (2015). Hochschulen für die demokratische Gesellschaft. In U. Haß (Hrsg.), *Kultur- und Medientheorie. Was ist eine Universität? Schlaglichter auf eine ruinierte Institution* (S. 27–34). transcript.

Verbeek, P.-P. (2005). *What things do: philosophical reflections on technology, agency, and design* (R. P. Crease, Übers.). The Pennsylvania State University Press.

Verbeek, P.-P. (2008). Cyborg intentionality: rethinking the phenomenology of human-technology relations. *Phenomenology and the Cognitive Sciences, 7*(3), 387–395. https://doi.org/10.1007/s11097-008-9099-x

Verbeek, P.-P. (2015). Beyond Interaction: A Short Introduction to Mediation Theory. *interactions*, *22*(3), 26–31. https://dl.acm.org/doi/epdf/10.1145/2751314

Yan, L. & Gašević, D. (2026). *Agentivism: a learning theory for the age of artificial intelligence.* https://doi.org/10.48550/arXiv.2604.07813

Teil II

Wahrheit, Verantwortung, Bildung

Die Universität als „Instanz der Wahrheit schlechthin“

Markus Stepanians

> „Leute wie Sie“, sagte der Berater [von US-Präsident George W. Bush], „gehören zu dem, was wir die ‚realitätsbasierte Gemeinschaft‘ nennen. Ihr glaubt, dass Lösungen aus dem sorgfältigen Studium der erkennbaren Realität hervorgehen. Aber so funktioniert die Welt heute nicht mehr“, fuhr er fort. „Wir sind jetzt ein Imperium, und wenn wir handeln, schaffen wir unsere eigene Realität. Und während ihr diese Realität untersucht – sorgfältig, wie ihr es eben tut –, werden wir erneut handeln und weitere neue Realitäten schaffen, die ihr ebenfalls untersuchen könnt, und so werden sich die Dinge entwickeln. Wir sind die Akteure der Geschichte … und ihr, ihr alle, werdet nur noch das untersuchen können, was wir tun.“
>
> R. Suskind ‘Faith, Certainty and the Presidency of George W. Bush’, *New York Times Magazine*, 17.10.2004

Wahrheit als zentraler Wert der Universität

Die Organisatorinnen unseres Symposiums haben uns aufgetragen, über die Idee der Universität im 21. Jahrhundert nachzudenken. Sie fragen unter anderem: „Was soll den Kern universitären Lehrens, Lernens und Forschens in Zukunft ausmachen?“ Und: „Welche Impulse können normative Leitbilder universitärer Bildung in klassischen Schriften (Humboldt, Newman, Jaspers) der Hochschulbildung heute geben?“ Die Antwort, die ich zur Diskussion stellen möchte, ist weit älter als die diesbezüglichen Überlegungen Newmans, Humboldts oder Jaspers‘. Sie hat ihre Wurzeln bei Platon und lautet: Der Kern allen Lehrens, Lernens und Forschens, also auch des universitären, war und ist das Streben nach Wahrheit. Wie sich in Krankenhäusern alles um Gesundheit dreht, Kunst und Theater sich dem Wert der Schönheit widmen und das Justizwesen auf Gerechtigkeit zielt, so sollten sich Universitäten als Institutionen verstehen, die in erster Linie dem Wert der Wahrheit verpflichtet sind. Mit den Worten des amerikanischen Philosophen Charles S. Peirce besteht echte Wissenschaft „in der sorgfältigen Suche nach der Wahrheit um der Wahrheit willen. […] Darin, den Bogen auf die Wahrheit zu richten, mit zielstrebigem Blick und kraftvollem Arm.“[1] Im Erfolgsfall treffen wir die Wahrheit und erkennen, was der Fall ist: die objektiven, unabhängig von uns existierenden Tatsachen, die in ihrer Summe die Welt konstituieren. Die zentrale Stellung, die Platon der Wahrheit in den Wissenschaften zuweist, ist kein europäisches Phänomen. Das zeigt sich in der Prominenz dieses Begriffs in den Leitsätzen zahlreicher Hochschuleinrichtungen überall auf der Welt. Hunderte von Universitäten definieren sich explizit im Rückgriff auf Wahrheit. Die alphabetische Liste reicht von der indischen Acharya Nagarjuna Universität (‚Alles gründet in Wahrheit‘) und der Bilkent Universität in der Türkei mit ihrem schlichten ‚Wahrheit‘ über die Cape Coast Universität in Ghana (‚Die Wahrheit ist unser Licht‘) und dem Churchill College in Cambridge, UK (‚Nicht mir musst Du glauben, sondern der Wahrheit‘), etc. pp. bis hin zur

[1] Zitiert nach S. Haack, „‘We Pragmatists’ … Peirce and Rorty in Conversation” in: Haack, S.: *Manifesto of a Passionate Moderate – Unfashionable Essays*, University of Chicago Press 1998. Soweit nicht anders angegeben sind alle Übersetzungen aus dem Englischen von mir.

Zhejiang Universität in China mit ihrem Wahlspruch ‚Die Wahrheit suchen und neue Wege beschreiten.'[2]

Natürlich schließt die Anerkennung der Wahrheit als Grundwert und Maßstab wissenschaftlichen Denkens andere Werte nicht aus. Fast ebenso häufig findest sich in den universitären Leitsätzen weltweit das Bekenntnis zu Wissen, Weisheit und Dienst am Gemeinwohl. Aber es scheint klar, dass diese und ähnliche Werte gegenüber dem Grundwert der Wahrheit derivativ sind und ihn voraussetzen. Denn Wissen und Weisheit sind ohne Wahrheit nicht zu haben, und der erfolgreiche Dienst am Gemeinwohl setzt Einsichten darüber voraus, worin es besteht. Dasselbe gilt für Hochschulen, die in ihren Wahlsprüchen den Wahrheitsbegriff nicht ausdrücklich erwähnen und stattdessen das Humboldt'sche Bildungsideal betonen. Beispielsweise sieht sich die Universität Hamburg, „der Forschung, der Lehre und der Bildung" verpflichtet.[3] Aber auch hier ist mit ‚Lehre' die Lehre von Wahrheiten, mit ‚Forschung' die Entdeckung von Wahrheiten und mit ‚Bildung' eine Form der Persönlichkeitsentfaltung gemeint, die eine Vielzahl wahrer Überzeugungen voraussetzt. In Wilhelm von Humboldts klassischer Definition besteht Bildung in der Realisierung des „wahren Zwecks des Menschen." Es mag unklar und umstritten sein, worin dieser „wahre Zweck" besteht. Klar ist jedoch, dass ein Leben in Irrtum, Täuschung und Illusion mit diesem Zweck unverträglich erscheint. Ein solider Grundstock wahrer Überzeugungen über die Welt ist eine unentbehrliche Voraussetzung für menschliches Überleben, ganz zu schweigen von einem guten menschlichen Leben. Ohne sie könnten wir unsere Lebenspläne nicht einmal im Ansatz erfolgreich umsetzen. Es ist daher nur folgerichtig, dass die von Humboldt gegründete Berliner Universität den begrifflichen Zusammenhang zwischen Bildung, Wissenschaft und Wahrheit in ihrem offiziellen Leitsatz ‚Bildung durch Wissenschaft' indirekt anerkennt. Eine explizite Formulierung desselben Gedankens könnte lauten: ‚Bildung durch das wissenschaftliche Streben nach Wahrheit'.

Die Überlegungen Newmans, Humboldts und Jaspers' scheinen mir kaum mehr als Erläuterungen dieses platonischen Grundgedankens der zentralen Bedeutsamkeit von Wahrheit zu sein. Die notwendige Orientierung an Wahrheit wird von diesen Autoren nur deshalb nicht eigens thematisiert, weil sie ihnen allzu selbstverständlich schien. Was könnte offensichtlicher sein, als dass universitäre Bildung nicht zu haben ist ohne Wahrheit und ohne das ernstliche Streben nach ihr, das wir Wahrhaftigkeit nennen? Aber in Krisenzeiten, wenn der Wahrheitsbegriff selbst unter Druck gerät und ernstlich in Frage gestellt wird (und mit ihm die Fundamentalunterscheidung zwischen wahr und falsch), kann es notwendig werden, sich dieser Selbstverständlichkeit wieder affirmativ und mit Nachdruck zu vergewissern. Eben dies tut Jaspers in seiner ‚Die Idee der Universität' von 1946. Die Schlussfolgerungen, zu denen sich Jaspers bezüglich der Bedeutsamkeit eines robusten Verständnisses von Wahrheit auch, aber nicht nur, aus politischen Gründen gedrängt sah, möchte

[2] https://en.wikipedia.org/wiki/List_of_university_and_college_mottos

[3] So jedenfalls steht es über dem Eingang des Hauptgebäudes der Universität Hamburg. Wie sich dieser Leitsatz mit den neuerdings obligatorischen ‚mission statements' verträgt, ist eine andere Frage. Im 21. Jahrhundert präsentiert sich die Universität Hamburg als ‚Flagship University: Innovating and Cooperating for a Sustainable Future in a Digital Age.' (https://www.uni-hamburg.de/uhh/doppeltransformation.html)

ich im Folgenden diskutieren. Beginnen werde ich jedoch zunächst mit George Orwell, der zur gleichen Zeit wie Jaspers die politischen Gefahren eines schwachen, falsch verstandenen Wahrheitsbegriffs mit bemerkenswerter Hellsichtigkeit gesehen hat.

„Der Begriff der objektiven Wahrheit verschwindet zunehmend aus der Welt"

Im Januar 1939 beginnt George Orwell seine Rezension eines neuen Buchs von Bertrand Russell mit der Bemerkung: „Wenn bestimmte Passagen in Bertrand Russells Buch *Power* ein wenig inhaltsleer scheinen, so liegt das allein daran, dass wir inzwischen an einem Tiefpunkt angelangt sind, an dem es die erste Pflicht intelligenter Menschen ist, das Offensichtliche erneut zu bekräftigen." Eine jener inhaltsleer wirkenden Passagen, in der Russell seiner Pflicht zur Betonung des Offensichtlichen nachkommt, ist diese: „Die Realität dessen, was von meinem eigenen Willen unabhängig ist, verkörpert sich für die Philosophie im Begriff der ‚Wahrheit'. Die Wahrheit meiner Überzeugungen hängt nach allgemeiner Auffassung in den meisten Fällen nicht von irgendetwas ab, das ich tun kann. […] W]enn ich glaube, dass Cäsar an den Iden des März ermordet wurde, liegt das, was meine Überzeugung wahr macht, gänzlich außerhalb meiner Willensmacht."[4]

Russell macht hier zwei Punkte. Angenommen, meine Überzeugung, dass Cäsar an den Iden des März ermordet wurde, wäre wahr. Dann würde erstens gelten, dass sie wahr ist aufgrund eines historischen Ereignisses, das am 15. März 44 v. u. Z. stattfand, nämlich ein politisch motivierter Mord. Dieser Mord ist „das, was meine Überzeugung wahr macht." Zweitens, wenn Cäsar an jenem fernen Märztag ermordet wurde, dann können wir diese Tat im Nachhinein weder ungeschehen machen noch irgendetwas daran ändern. Wir mögen Cäsars Ermordung bedauern oder begrüßen. Aber der Wahrmacher der Überzeugung, das historische Ereignis, wird durch unser heutiges Denken und Fühlen nicht verändert. Natürlich ist es richtig, dass wir uns auf dieses Ereignis durch Begriffe beziehen, die eine Geschichte haben, kulturellen Einflüssen unterworfen sind und oft Bewertungen implizieren. Es sind unsere Begriffe, wie wir sie heute verstehen – wessen sonst? Für *uns* macht es einen Unterschied, ob wir jenes Ereignis als einen heimtückischen Mord, einen Tyrannenmord oder – betont wertneutral – als Tötung beschreiben. Die Begriffe, die den Inhalt meiner Überzeugung bilden, bestimmen, unter welchen Bedingungen sie wahr ist. Ob diese Wahrheitsbedingungen jedoch erfüllt sind, hängt allein von der Welt ab. Alles Fürwahrhalten ist – der Intention nach – das Anerkennen eines Wahrseins, das sich um unsere (voluntativen oder kognitiven) mentalen Einstellungen zu ihm nicht schert. Wenn eine Überzeugung wahr ist, dann ist sie es aufgrund der Art und Weise, wie die weitgehend unabhängig von uns existierende Welt ist.

All das mag so trivial sein, dass es kaum der Erwähnung wert erscheint. Aber Orwell glaubt in einer Zeit zu leben, in der diese und andere Selbstverständlichkeiten ernstlich in Gefahr sind. Die gebetsmühlenartige Wiederholung trivialer Wahrheiten ist in Orwells Augen die

[4] B. Russell, *Power – A New Social Analysis*, Routledge 2004 (Orig. 1938), S. 210. Orwells Rezension erschien im Januar 1939 in der Literaturzeitschrift *The Adelphi*.

erste Pflicht intelligenter Menschen, weil diese Trivialitäten die Existenz einer festen, weitgehend unveränderlichen Welt verbürgen, die gegen politische Manipulation und Propaganda immun ist. In einem Essay von 1943 schreibt Orwell: „Der Begriff der objektiven Wahrheit verschwindet zunehmend aus der Welt […]. Der Nationalsozialismus bestreitet ja ausdrücklich, dass es so etwas wie ‚die Wahrheit' überhaupt gibt. Es gibt zum Beispiel keine ‚Wissenschaft' im eigentlichen Sinne. Es gibt nur ‚deutsche Wissenschaft', ‚jüdische Wissenschaft' usw. Das implizite Ziel dieser Denkweise ist eine albtraumhafte Welt, in der der Führer oder eine herrschende Clique nicht nur die Zukunft, sondern auch *die Vergangenheit* kontrolliert. Wenn der Führer über dieses oder jenes Ereignis sagt: ‚Das ist nie geschehen' – nun, dann ist es nie geschehen. Wenn er sagt, dass zwei plus zwei fünf ist – nun, dann ist zwei plus zwei fünf. Diese Aussicht macht mir sehr viel mehr Angst als Bomben – und nach unseren Erfahrungen der letzten Jahre ist das keine leichtfertige Behauptung."[5] Gegen den Albtraum einer „sich ständig wandelnden, phantasmagorischen Welt, in der Schwarz morgen schon Weiß sein kann und das Wetter von gestern per Dekret geändert werden kann", so Orwell weiter, helfe nur die unerschütterliche Gewissheit, dass, „so sehr die Wahrheit auch verleugnet wird, die Wahrheit dennoch, sozusagen hinter unserem Rücken, weiterexistiert."[6] In Orwells dystopischem Roman *1984* klammert sich der tragische Held Winston Smith verzweifelt an die Gewissheiten, die das allzu Offensichtliche in einem politischen Regime gewährt, dessen totalitärer Anspruch vor nichts haltmacht. Die von Winston aufgezählten Trivialitäten sind ihm Garanten und Fixpunkte für die Existenz einer Welt, die ungeachtet der systematischen Lügen ‚der Partei', also gewissermaßen hinter ihrem Rücken, weiterexistiert. Die Widerspenstigkeit und Hartnäckigkeit der Tatsachen sind in Winstons Welt das einzige, das der überwältigenden Macht der Partei Grenzen setzt. Mit Blick auf diese Gewissheiten sieht sich Winston im Recht, wenn er gegen die Partei aufbegehrt und auf triviale Alltagswahrheiten als unverrückbare Grenzsteine der Macht der Partei verweist: „Die Partei lehrt einem, der Evidenz unserer Augen und Ohren nicht zu trauen. Das war ihr entscheidendes, wichtigstes Gebot. […] Und dennoch war er im Recht! Sie hatten unrecht und er hatte recht. Das Offensichtliche, das Törichte und das Wahre mussten verteidigt werden. Binsenwahrheiten sind wahr, haltet euch daran fest! Die feste Welt existiert, ihre Gesetze ändern sich nicht. Steine sind hart, Wasser ist nass, Gegenstände, die nicht aufgehalten werden, fallen zum Erdmittelpunkt. Mit dem Gefühl …, dass

[5] G. Orwell, 'Looking Back on the Spanish War', S. 258 f.; in: *The Collected Essays, Journalism and Letters of George Orwell*, Vol. II, 1968. Es ließe sich schnell zeigen, dass Orwells knappe Charakterisierung die Wissenschafts- und Wahrheitsauffassung der Nationalsozialisten den Nagel auf den Kopf trifft. *Mutatis mutandis* gilt das nach Orwell auch für Stalins UdSSR.

[6] Ibid, S. 258 f.; in: *The Collected Essays, Journalism and Letters of George Orwell*, Vol. II, 1968; meine Übersetzung. Obwohl Orwell seine These vom Verschwinden des Begriffs einer objektiven Wahrheit „aus der Welt" hier unter Verweis auf Hitler-Deutschland illustriert und Stalins Sowjetunion das Modell der Propagandamaschinerie ‚der Partei' in *1984* ist, meint er tatsächlich „die Welt" einschließlich der westlichen Demokratien. Orwell reflektiert in diesem Essay seine Erfahrungen im spanischen Bürgerkrieg und wie über diesen Krieg in der Presse, auch der englischen, berichtet wird: „Ich hatte schon früh im Leben bemerkt, dass die Zeitungen über kein Ereignis jemals wahrheitsgemäß berichten, aber in Spanien sah ich zum ersten Mal Zeitungsberichte, die in keinerlei Beziehung zur Wirklichkeit standen, nicht einmal der, die eine gewöhnliche Lüge impliziert. Dort wurde über große Schlachten berichtet, wo gar nicht gekämpft wurde und die Tötung von Hunderten mit Schweigen übergangen" (S. 256).

er ein wichtiges Axiom darlegt, schrieb er: Freiheit ist die Freiheit zu sagen, dass zwei plus zwei vier ist. Wenn das zugestanden wird, folgt alles weitere."[7]

Karl Jaspers' ‚Idee der Universität' in der Neufassung von 1946

Schon 1923 hatte Jaspers ein kleines Büchlein mit dem Titel ‚Die Idee der Universität' publiziert. 1946 legt er unter der gleichen Überschrift eine Überarbeitung vor, die mit dem Werk von 1923 kaum mehr etwas gemein hat. Im Vorwort vom Mai 1945 rechtfertigt Jaspers die Neufassung mit der „moralischen Vernichtung der Universität" der letzten zwölf Jahre. Der vielleicht auffälligste Unterschied zwischen dem Text von 1923 und dem von 1946 betrifft die Rolle der Wahrheit in den Wissenschaften. Beide Texte fassen die Universität als eine Institution, die der Wissenschaft und dem Prozess des Erkennens gewidmet ist. Aber der ältere Text von 1923 betont den *Prozess* des Erkennens, seinen prekären, von Ungewissheit geprägten Charakter und behauptet seine prinzipielle Unabgeschlossenheit. Auf Wahrheit als Ziel dieses Erkenntnisprozesses kommt Jaspers 1923 ausdrücklich erst spät zu sprechen, und als er es schließlich tut, geschieht es gewissermaßen mit spitzen Fingern. Der frühe Jaspers setzt ‚Wahrheit' in An(gst)führungszeichen[8] und bekundet im alten Streit zwischen ‚Dogmatikern' und ‚Skeptikern' seine Sympathien für letztere. Die philosophische Skepsis habe gezeigt, so der Jaspers von 1923, dass Wahrheit in einem „endgültigen" oder „absoluten" Sinne illusorisch sei. Der Erkenntnisprozess gelange nie an sein Ziel. Er bestehe „immer nur in der Bewegung des Erkennens, nicht im festen Besitz erworbenen, endgültigen Wissens, sie hat nie die endgültige Wahrheit […] Das dogmatische Bewusstsein, als ob die absolute Wahrheit greifbar und in Besitz wäre, wird durch die Kräfte der Skepsis zerstört."[9]

Im Gegensatz dazu ist in der Neufassung von 1946 von Skepsis gegenüber Wahrheit und der Möglichkeit erfolgreicher Wahrheitsfindung nichts mehr zu spüren. Gleich auf der ersten Seite bestimmt Jaspers die Universität als Ort „bedingungsloser Wahrheitsforschung", und er lässt keinen Zweifel daran, dass Wahrheitsforschung nicht nur erfolgreich sein kann, sondern es oft auch ist. Die Universität sei „eine Anstalt mit realen Zielen, die jedoch erreicht werden," nämlich die Suche nach Wahrheit in der Forschung und die Weitergabe der gefundenen Wahrheiten in der Lehre: „An der Universität sind Menschen vereinigt in einer Institution zu dem Beruf, die Wahrheit durch Wissenschaft sowohl zu *suchen* als auch zu *überliefern*." Den Skeptikern und seinem früheren Selbst von 1923 hält Jaspers nun entgegen, dass diese doppelte Aufgabe auch dann zu bewältigen sei, wenn wir keine klare Vorstellung davon haben, „was Wahrheit ist und was diese Wahrheitsbemächtigung."[10]

---

[7] G. Orwell, *1984*, London 1949. Am Ende des Romans wird Winston auf grausame Weise klargemacht, dass selbst eine mathematische Wahrheit wie ‚2+2=4' in einem totalitären Regime kein unumstößliches Axiom ist, sondern bestenfalls Ausdruck einer Hoffnung, die unweigerlich enttäuscht wird. Am Ende muss Winston, wie Orwell in seiner Russell-Rezension sagt, „einsehen, dass ‚Wahrheit ist groß und wird sich durchsetzen' ein Gebet ist und kein Axiom."

[8] K. Jaspers, *Schriften zur Universitätsidee*, hrsg. von O. Immel, Basel 2016. (*Karl Jaspers Gesamtausgabe* / Abt. 1, Werke; Bd. 21), S. 11.

[9] Ibid. S. 15. – In der Philosophie ist ‚dogmatisch' seit dem 1. Jahrhundert u. Z. der Gegenbegriff zu ‚skeptisch'.

[10] K. Jaspers, *Die Idee der Universität* (1946), S. 109.

Natürlich bestreitet Jaspers auch 1946 nicht, dass das menschliche Erkenntnisvermögen begrenzt, prekär und fallibel ist. Aber diese epistemischen Defizite zeigen weder, dass Wahrheit nicht existiert, noch, dass wir sie nie erkennen können.

Es ist nicht klar, was genau Jaspers mit „endgültigen" oder „absoluten" Wahrheiten meint. Sollten damit jedoch Wahrheiten gemeint sein, die unter keinen Umständen in Frage gestellt oder revidiert werden können, dann kann auch ein ‚Dogmatiker' gelassen einräumen, dass es sie nicht gibt. Anfang der 1950er betont der amerikanische Philosoph Willard van Orman Quine in einem berühmten Aufsatz, dass „keine Aussage gegen Revision immun ist."[11] Gleichwohl hält er daran fest, dass wir viele Wahrheiten mit hinreichender Sicherheit wissen. In Quines Augen pocht Winston völlig zurecht darauf, dass Steine hart sind und Wasser nass, dass Gegenstände, die nicht aufgehalten werden, zum Erdmittelpunkt fallen und dass 2 + 2 = 4 ist. Es mögen Umstände denkbar sein, unter denen auch diese Gewissheiten ernstlich hinterfragt werden müssen. Aber selbst dann wäre der Zweifel nur verständlich vor einem breiten Hintergrund anderer anerkannter Wahrheiten. Shakespeares' Macbeth mag in einer konkreten Situation seinen Sinnen mit guten Gründen misstrauen und sich ernstlich fragen, ob das, was er dem Anschein nach vor sich sieht, wirklich ein Dolch ist. Würde Macbeth jedoch seinen Zweifel auf *alle* Alltagsgewissheiten ausdehnen, dann würden wir uns mit Recht fragen, ob er noch bei Sinnen ist.[12] Grundsätzlich scheint jede Überzeugung für sich genommen revidierbar. Es ist uns jedoch unmöglich, in skeptischem Übermut das gesamte Netzwerk unserer Überzeugungen auf einen Schlag zu verwerfen. In dieser Hinsicht sind wir, wie der österreichische Philosoph Otto Neurath zu Beginn der 1930er in einem ähnlichen Kontext formuliert, „wie Schiffer, … die ihr Schiff auf hoher See umbauen müssen, ohne es jemals in einem Dock zerlegen und aus besten Bestandteilen nur errichten zu können."[13]

Eine zweite These, die Jaspers schon 1923 formuliert und 1946 mit Nachdruck wiederholt, lautet, dass die Wahrheitssuche keiner Zweckbindung unterworfen sein darf. Jaspers wendet sich erneut gegen die „Subalternisierung der Wissenschaft durch Unterordnung unter die Zwecke der Technik und Lebenspraxis."[14] Die Brauchbarkeit und lebenspraktische Verwertbarkeit der gefundenen Wahrheiten seien, wenn es sie denn gäbe, nur ein Nebeneffekt, wenn auch ein willkommener. Letztlich, so Jaspers, gehe es immer um die Erforschung der Wahrheit um der Wahrheit willen. In diesem Sinne ist die Universität für Jaspers eine „Instanz der Wahrheit *schlechthin*."[15]

[11] W. Quine, *Two Dogmas of Empiricism*, The Philosophical Review 60, 1951: 20-43, § VI.
[12] „Wer an allem zweifeln wollte, der würde auch nicht bis zum Zweifel kommen. Das Spiel des Zweifelns selbst setzt schon die Gewissheit voraus." L. Wittgenstein, *Über Gewissheit*, Suhrkamp 1984, §115.
[13] O. Neurath, ‚Protokollsätze', Erkenntnis 3, S. 204–214 (1932), S. 206.
[14] K. Jaspers, *Die Idee der Universität* (1946), S. 117.
[15] K. Jaspers, *Philosophische Autobiographie*, München 1977, S. 57; meine Hervorh.

Wie KI die Universität als “Instanz der Wahrheit” beeinflusst – ein Kommentar zu Markus Stepanians

Gabi Reinmann & Frank Vohle

Die Universität als „Instanz der Wahrheit schlechthin“. Markus Stepanians bestimmt in seinem Essay mit diesem Titel Wahrheit als den unhintergehbaren Grundwert der Universität. Gemeint ist damit nicht ein Wert unter vielen, sondern derjenige, von dem andere akademische Werte abhängen, wie sie in aktuellen universitären Leitbildern oft genannt werden, etwa das Bekenntnis zu Wissen, Weisheit und Dienst am Gemeinwohl. Exemplarisch, nämlich unter Rückgriff auf Karl Jaspers und George Orwell, erörtert Stepanians, wie fragil dieser Grundwert in Krisenzeiten werden kann: Besonders eindrucksvoll zeigt sich das in Jaspers´ beiden Fassungen der Schrift zur „Idee der Universität“. Während er 1923, in der ersten Fassung, die Vorläufigkeit und Ungewissheit des Erkennens betont, insistiert er 1946, nach den Erfahrungen mit dem Nationalsozialismus, auf einer bedingungslosen Wahrheitsforschung ohne Zweckbindung; Erkenntnisgrenzen und menschliche Fehlbarkeit, die anzuerkennen seien, würden weder die Existenz von Wahrheit noch die grundsätzliche Möglichkeit, sie zu erkennen, widerlegen. Die politischen Risiken einer schwachen oder falsch verstandenen Wahrheitsauffassung hat im gleichen Zeitraum auch George Orwell erkannt und literarisch aufgearbeitet. In der Figur Winston Smiths im Roman 1984 wird illustriert, wie nur mehr das Festhalten an einfachen, offensichtlichen Tatsachen dem Anspruch einer allmächtigen Partei – im Prinzip – noch Grenzen setzen kann. Die Wurzeln von Wahrheit als Grundwert in dem Sinne, wie sie bei Orwell und Jaspers für die Wissenschaft deutlich werden, führt Stepanians auf Platon zurück, sind also weitaus älter als Schriften zur Universität. Mit Jaspers und Orwell veranschaulicht Stepanians vor allem, inwiefern das scheinbar selbstverständliche Streben nach Wahrheit als Kern der Universität in Krisenzeiten unter Druck geraten kann und dann offensiv verteidigt werden muss. Auslöser für die Verteidigung der Wahrheit sind bei seinen beiden Protagonisten die verheerenden Folgen des Nationalsozialismus.

Totalitarismus als Bedrohung für die Wahrheit. Stepanians Essay konzentriert sich auf die politisch-historische Krise (oder besser: Katastrophe) infolge totalitärer Herrschaft vor und im zweiten Weltkrieg und arbeitet heraus, wie eine solche nicht nur den Begriff, sondern auch die Existenz von Wahrheit geradezu ausradieren kann. Die Universität als Institution, so zitiert Stepanians Karl Jaspers´ Rechtfertigung der Neufassung seiner Schrift zur Idee der Universität, sei in dieser Zeit moralisch vernichtet worden, hatte es also nicht geschafft, die Wahrheit als Grundwert in der Krise wirksam zu verteidigen. Nun mag es auf den ersten Blick weit weg erscheinen, die angeführten historischen und literarischen Beispiele in Stepanians Essay mit Künstlicher Intelligenz (KI) an Universitäten in einen Zusammenhang zu bringen. Doch eben dies, nämlich die Identifikation totalitärer Momente in der Entwicklung und Ausbreitung von KI ist in den letzten Jahren Gegenstand wissenschaftlicher

Reflexionen geworden[1]: So analysiert Alice Watanabe (2023) Anwendungen von KI im Hochschulbereich unter Rückgriff auf Hannah Arendt daraufhin, welche Parallelen es zwischen KI-Technologien und totalitären Strukturen geben kann. Rainer Mühhoff (2025) arbeitet heraus, inwieweit die Funktionsweise von KI Parallelen zu autoritären Ideologien hat; darüber hinaus warnt er davor, dass KI in den Händen weniger Tech-Konzerne ein Werkzeug für totale Machtanhäufung sein kann. Mark Coeckelbergh (2026) demonstriert, inwiefern die Digitalisierung, insbesondere soziale Medien und KI, ökonomischen wie auch politischen Interessen dienen und dabei autoritäre und totalitäre Fantasien beflügeln. So betrachtet ließe sich das gerade begonnene KI-Zeitalter ebenfalls als Krise für den Grundwert Wahrheit deuten – nämlich über den Weg der Begünstigung totalitären Denkens und Handelns. Aber wir sehen noch einen direkteren, besonders zerstörerischen, „Schleichweg" für die KI in den Prozess der Wahrheitsfindung an Universitäten: Sie nimmt Einfluss auf die Bedeutung von Wahrheit ebenso wie auf die Bedingungen für das Finden, Prüfen und Aushandeln wahrer Erkenntnisse.

Der Bedeutungsraum von Wahrheit. Was unter Wahrheit zu verstehen ist, wird in Stepanians Essay zwar nicht eingehend behandelt, aber doch angedeutet: In den argumentativ eingewobenen Beispielen klingt vor allem ein naturwissenschaftlicher Wahrheitsbegriff im Sinne einer Tatsachen- oder Faktenfeststellung an, wie er in der neuzeitlichen, also empirischen, Wissenschaft üblich ist. Andere Wahrheitserfahrungen, die sich dieser Methodik entziehen, etwa die der Kunst oder des hermeneutischen Verstehens, sind damit weitgehend ausgeblendet[2]. Eine nochmals andere Position zur Wahrheit, die man hier ergänzen könnte (und sollte), nehmen Vertreter des (amerikanischen) Pragmatismus wie Charles Sanders Peirce, William James und John Dewey[3] ein: Sie rücken – bei unterschiedlicher Nähe zum Realismus – von einem reinen Korrespondenzgedanken ab und binden Wahrheit stattdessen an so etwas wie Bewährung im Forschungs- oder Handlungsprozess oder an den praktischen Nutzen einer Idee. In späteren Schriften zum Konstruktivismus wird diese Denkweise im Begriff der Viabilität noch radikaler postuliert: Eine Erkenntnis gilt nicht, weil sie die Realität abbildet, sondern weil sie sich im Handeln bewährt, ohne mit der Welt in Widerspruch zu geraten. Auch in der Systemtheorie Luhmanns (1999) spielt der Abgleich mit der Wirklichkeit keine Rolle: Da jede Prüfung selbst nur Beobachtung und kein direkter Weltzugriff sei, lasse sich nie unabhängig prüfen, ob eine Aussage der Wirklichkeit entspricht. Trotzdem macht Niklas Luhmann die Unterscheidung von wahr und unwahr zum Leitcode der Wissenschaft bzw. der Universität; Wahrheit ist hier aber keine Eigenschaft, sondern ein Werkzeug der Kommunikation. Es gibt also, so kann man in aller Kürze festhalten, seit Jahrzehnten unterschiedliche, teils konkurrierende Versuche, Wahrheit zu bestimmen; der Bedeutungsraum für diesen Begriff ist größer als man zunächst denken mag. Für die Universität als Instanz für Wahrheit sollte dies keineswegs ein Problem sein – im Gegenteil: Das Ringen um Wahrheit und Wahrheitsbegriffe dürfte eine

[1] Es sei betont, dass hier Totalitarismus in der NS-Zeit nicht mit dem gleichgesetzt werden soll, das in der KI-Zeit beobachtbar ist; vielmehr geht es den zitierten Autoren darum, autoritäre bis totalitäre Momente oder Mechanismen zu erkennen.

[2] Dies ist die Argumentation vor allem von Hans Georg Gadamer (vgl. Figal, 2011)

[3] Einen einfachen, aber instruktiven Überblick bietet Duru (2017)

mindestens ebenso wichtige Rolle spielen, wie das Finden von Wahrheit, wenn es um Wahrheit als *Grundwert* geht. Eben dieses Ringen in einem Bedeutungsraum von Wahrheit, der eine gewisse Varianz zulässt, könnte infolge von KI ins Wanken geraten.

Neue Wahrheitskriterien durch KI. KI nimmt derzeit unmittelbar Einfluss auf die epistemische Praxis und darauf, was Wahrheit bedeutet. In aktuellen Texten zu KI in der Wissenschaft wird dargelegt, inwieweit inzwischen von einer künstlichen, synthetischen oder algorithmischen Wahrheit auszugehen ist: So sieht etwa Palese (2026) das klassische Korrespondenzkriterium (Wahrheit als Übereinstimmung mit der Realität) durch ein neues Kriterium ersetzt, nämlich: rechnerische Plausibilität und Plattformintegration, und das heißt: Entscheidend ist jetzt, was plausibel klingt und zur Logik von digitalen Plattformen passt; es entsteht eine „künstliche Wahrheit". Damit verschiebt sich auch die epistemische Autorität – weg von Fachleuten und Institutionen (wie der Universität) hin zu augenscheinlicher Kohärenz und Passung in KI-Systemen. Shin (2026) ergänzt, dass mit dem Einsatz von KI in der Forschung Daten und Erkenntnisse nicht mehr nur von Menschen geprüft, sondern von KI validiert werden. Das hat zur Folge, dass KI-Systeme selbst festlegen, was als „wahr" gelten kann. Diese Systeme aber sind nicht neutral, sondern bergen (wie oben bereits erwähnt) versteckte Annahmen, externe Interessen und Verzerrungen in sich. Wer die Systeme baut, bestimmt mit, was Wahrheit ist; diese avanciert zu einer algorithmischen Wahrheit. Zusammengenommen ließe sich festhalten: An die Stelle von Korrespondenz mit der Realität tritt rechnerische Plausibilität; die epistemische Autorität auf menschlicher Seite wird reduziert, indem KI zur Prüfinstanz wird; wissenschaftliche Verantwortung verschiebt sich von Institutionen auf Plattformen. Algorithmische oder künstliche Wahrheit in diesem Sinne ist weder mit dem neuzeitlichen Wahrheitsbegriff der Fakten- oder Tatsachenfeststellung noch mit anderen menschlichen Wahrheitserfahrungen vereinbar. Man mag eine oberflächliche Nähe zur Viabilität in dem Sinne ausmachen, dass sich auch rechnerische Plausibilität an ihrem „Funktionieren" bemisst. Doch anders als Viabilität im Konstruktivismus oder Bewährung in der Welt des Handelns im Pragmatismus ist sie eben nicht an gelebte Erfahrung und Weltbezug rückgebunden, sondern allein an Optimierung und Effizienz der „Wahrheit" erzeugenden KI-Systeme.

Strukturelle Vorbereitung durch die Wissenschaft selbst. Es greift allerdings zu kurz, allein in der Ausbreitung und Entwicklung von KI die Ursache für einen Wandel der epistemischen Praxis zu sehen oder zu suchen. Wettbewerbsdruck und Effizienzstreben leiten den Universitätsalltag schon länger und haben den Boden für KI in der Wissenschaft bereitet, oder anders und radikaler formuliert: KI vollzieht, was strukturell längst angelegt ist. So führt etwa Manabu (2026) aus, dass das gängige Prinzip „Publish-or-Perish" (sinngemäß: wer nicht viel veröffentlicht, hat keine Karrierechancen in der Wissenschaft) durch KI keineswegs gestört, sondern vielmehr vollendet werde: Das ohnehin schon lange durch Metriken (wie Impact Faktoren) getriebene Publikationssystem werde nun durch Maschinen konsequent zu Ende gespielt, denn: Diese beherrschen das wissenschaftliche Publizieren nicht nur genauso gut, sondern (formal) noch besser als Menschen. Ergänzend und passend dazu argumentieren Lindebaum et al. (2025), dass sich Menschen mit dem Einzug von KI in die Wissenschaft tendenziell von epistemischen Akteuren zu epistemischen Konsu-

menten wandeln; auf diese Weise werde organisierte Unreife oder Unmündigkeit begünstigt: Diese Entwicklung, so Lindebaum et al. (2025), untergräbt das Humboldt'sche Ideal der Hochschulbildung als fortschrittliches Kulturprojekt und die intellektuelle Entfaltung der Universitätsmitglieder. Zur Verschiebung des Wahrheitskriteriums infolge von KI kommt, so ließen sich die beiden Beiträge zusammenfassen, eine von der Wissenschaft selbst verursachte institutionell-strukturelle Erosion der Bedingungen für das Finden von und Ringen um Wahrheit. Dies schwächt unmittelbar auch die Universität als Wahrheitsinstanz, was sicher kein bloß akademisches Problem ist, sondern ein Verlust an menschlicher Selbstbestimmung bedeutet.

Reibung zur Wahrheitsfindung. Ähnlich wie die zitierten Autoren zuvor führt Nowotny (2025) aus, wie KI menschlichen Aufwand und menschliche Fehler in Forschungsprozessen minimiert sowie die Effizienz in einem bislang nicht gekannten Ausmaß nach oben treibt; das könne man vor allem in den Naturwissenschaften an vielen Beispielen (z.B. Proteinfaltung) sehen – dort also, wo auch das klassische Wahrheitsverständnis zu verorten ist. Zudem verändere KI nicht nur die Forschungsorganisation, sondern ebenso die Forschungsmethode, und das in wohl allen Disziplinen. Besonders hebt Nowotny (2025) aber die Chancen von KI für größere interdisziplinäre Teams hervor: Eine entscheidende Folge des Einsatzes von KI sieht sie hier darin, dass bestehende Disziplingrenzen leichter überwunden und wissenschaftliche Domänen problemloser kombiniert werden könnten. Zugleich aber sei genau dies mit enormen Risiken verbunden: Eine KI-basierte Interdisziplinarität macht überflüssig, dass Experten geduldig die Fachsprache oder Bedeutung bestimmter Konzepte oder Methoden des jeweils anderen erlernen; KI ermöglicht eine Zusammenarbeit ohne Bemühen um gegenseitiges Verständnis und eröffnet schnellere Wege, Konsens und Resultate zu erzielen. Meinungsverschiedenheiten und Langsamkeit menschlicher Kommunikation werden überwunden, bestehende Unterschiede kalibriert und ausgeglichen, oder kurz: KI verspricht eine reibungslose Welt[4]. Für die Universität als Wahrheitsinstanz aber sollte Reibung, verstanden als Abwägung, epistemische Vielfalt und Konflikt im Sinne einer produktiven Kraft, essenziell sein – sofern man keine künstliche oder algorithmische Wahrheit anstrebt. Nun gibt es Autoren wie Cruz-Aguilar (2025), die in diesem Zusammenhang vorschlagen, die menschliche epistemische Praxis zwar nicht zu ersetzen, aber zu rekonfigurieren und auf eine hybride Mensch-Maschine-Praxis der Wahrheitsfindung zuzuschneiden. Nowotny (2025) hingegen fordert dazu auf, Reibung bewusst zu kultivieren und die Universität (wieder) zu einem Ort zu machen, an dem Wahrheit – auch eine gewisse Pluralität von Wahrheit im oben skizzierten Sinne – in mühsamer, aber genuin menschlicher und dialogischer Auseinandersetzung gesucht, geprüft und – je nach Disziplin – auch ausgehandelt wird.

Im Dialog zur Wahrheit. Diese Forderung nach einer Kultivierung von Reibung erinnert an die wohl älteste Gestalt europäischer Wahrheitssuche in der Antike, auf die auch Stepanians in seinem Essay verweist: den sokratischen Dialog. Sokrates zielte in seinen berühmten Dialogen gerade nicht auf schnelle Einigung ab, sondern verharrte bewusst im Zustand des Nicht-mehr-weiter-Wissens (Aporie) als einer notwendigen Phase im Erkenntnis-

[4] Siehe dazu auch den Beitrag von Judith Simon in diesem Essay-Band

prozess (z.B. Callard, 2025): Es ging Sokrates nicht um eine effiziente Wahrheitsfindung, sondern darum, bestehende Überzeugungen in Frage zu stellen, zu widerlegen (oder selbst widerlegt zu werden) und Widersprüche aufzuspüren – quasi eine philosophische Urform dessen, was Nowotny (2025) Reibung nennt. Darüber hinaus betont Callard (2025), dass für Sokrates ein Philosophieren und die Suche nach Wahrheit nur dialogisch, also mit einem menschlichen Gegenüber, möglich seien. Wahrheitssuche ist dialogisch[5] und steht damit in einem deutlichen Kontrast zu einer künstlichen Wahrheit, die sich in selbstbezüglichen KI-Systemen findet, in denen Kohärenz und Plausibilität die bestimmenden Wahrheitskriterien sind. Diese Perspektive wird in den uns bekannten internationalen Auseinandersetzungen zu KI in der Wissenschaft bislang kaum beleuchtet und stärkt Stepanians Grundwert der Wahrheit, deren Wurzeln er selbst bei Platon verortet.

Zwei Formen der Bedrohung. Stepanians Rückgriff auf Jaspers und Orwell zeigt, wie totalitäre Herrschaft den Grundwert Wahrheit an der Universität von außen bedroht – durch Zwang und Zerstörung einer bis dahin bestehenden wissenschaftlichen Kultur. Watanabe, Mühlhoff und Coeckelbergh arbeiten heraus, inwieweit sich in heutigen KI-Systemen und deren Nutzung ebenfalls totalitäre Momente finden lassen. Internationale Autoren, wie wir sie hier exemplarisch eingebunden haben, beschreiben Mechanismen, welche die Universität als Wahrheitsinstanz auf eine andere, leisere, aber nicht weniger wirkungsvolle Weise bedrohen: Sie löschen keine Fakten aus, verschieben aber die Kriterien, nach denen entschieden wird, was als wahr gilt. Jaspers musste die Idee der Universität 1946 unter dem Eindruck einer sichtbaren Katastrophe neu fassen. Aus unserer Sicht müsste sich die Universität angesichts von KI einer solchen Herausforderung ebenfalls stellen. Ein Anker dafür könnten Reibung im Sinne von Helga Nowotny und Judith Simon sowie der (sokratische) Dialog im Sinne zum Beispiel von Agnes Callards sein. Beides sind Prinzipien, um die sich Universitäten kümmern könnten und sollten, um den Grundwert Wahrheit zu erhalten oder wieder herzustellen und sich einer Etablierung von ökonomisch motivierten Wahrheitsbegriffen, die sich künstlich oder algorithmisch nennen, zu widersetzen.

Literatur

Callard, A. (2025). *Open Socrates: The case for a philosophical life*. New York: W. W. Norton & Company.

Coeckelbergh, M. (2026). Technofascism: AI, Big Tech, and the new authoritarianism. *AI & Society*, 41, 5163-5176. https://doi.org/10.1007/s00146-026-02862-9

Cruz-Aguilar, M.A. (2026). The epistemic revolution of AI: reconfiguring the foundations of scientific knowledge. *AI & Socieyt,* 41, 2041-2057 (2026). https://doi.org/10.1007/s00146-025-02658-3

Duru, M. (2017). Es war einmal in Amerika – die Story des Pragmatismus. *Philosophie Magazin*. https://www.philomag.de/artikel/es-war-einmal-amerika-die-story-des-pragmatismus

[5] Gemeint ist der Dialog zwischen Menschen

Feige, D.M. & Pollmann, A. (2026). Digitale Entgeisterung. *Philosophie Magazin*. https://www.philomag.de/artikel/digitale-entgeisterung

Figal, G. (Hrsg.) (2011). Hans-Georg Gadamer: Wahrheit und Methode

Lindebaum, D., Nolan, E., Ashraf, M., Islam, G. & Ramirez, M. (2025). The transformation of epistemic agency and governance in higher education through Large Language Models: toward a future of organized immaturity. *Organization Studies*. https://doi.org/10.1177/01708406251392002

Luhmann, Niklas (1990): *Die Wissenschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.

Manabu, T. (2026). The machine consumes itself: Artificial intelligence and the end of publish-or-perish. *Postdigital Science and Education*, 1-24. https://doi.org/10.1007/s42438-026-00666-0

Mühlhoff. R. (2025). *Künstliche Intelligenz und der neue Faschismus*. Ditzingen: Reclam.

Nowotny H. (2025). Bringing back friction: Reflections on a humanistic culture of AI interdisciplinarity. *Cambridge Forum on AI: Culture and Society*, 1, 1-12. doi:10.1017/cfc.2025.10005

Palese, R. (2026). Artificial truth: Algorithmic power, epistemic authority, and the crisis of democratic knowledge. *Societies*, *16*(3), 102. https://doi.org/10.3390/soc16030102

Shin, D. (2026). Automating epistemology: how AI reconfigures truth, authority, and verification. *AI & Society*, 41, 1553-1559. https://doi.org/10.1007/s00146-025-02560-y

Watanabe, A. (2023). Exploring totalitarian elements of artificial intelligence in higher education with Hannah Arendt. *International Journal of Technoethics*, 14(1), 1-15. http://doi.org/10.4018/IJT.329239

Wahrheit und der explanatorische Primat der Welt – eine Replik auf den Kommentar von Gabi Reinmann und Frank Vohle

Markus Stepanians

Der Kommentar von Gabi Reinmann und Frank Vohle zu meinem Symposiumsbeitrag ist fair und angemessen. In zwei Punkten fühle ich mich gleichwohl missverstanden. *Erstens* schreiben sie mir die These zu, dass der politische Totalitarismus eine „Bedrohung für die Wahrheit" darstellt und „nicht nur den Begriff, sondern auch die Existenz von Wahrheit geradezu ausradieren kann." Genau das glaube ich nicht. Keine Macht der Welt kann die Existenz von Wahrheit vernichten, weil sie die Welt als solche nicht vernichten kann. Auch unser impliziter Begriff von Wahrheit kann durch verfehlte Theorien und Manipulationen nur verzerrt und verfälscht, aber nicht „ausradiert" werden. Er kann nicht vernichtet werden, weil er für unser Selbstverständnis als (mehr oder weniger) rationale Wesen konstitutiv ist. *Zweitens* suggerieren Reinmann und Vohle, dass das von mir favorisierte objektive und robuste Wahrheitsverständnis eines sein müsse, das Wahrsein als Übereinstimmung (‚Korrespondenz') mit den Tatsachen konzipiert. Besonders mit Blick auf den Jasperschen Slogan von der ‚Universität als Instanz der Wahrheit schlechthin' plädieren sie stattdessen für eine „gewisse Varianz". Für Reinmann und Vohle ist der Slogan nur akzeptabel, wenn ‚Wahrheit' hier nicht in einem engen korrespondenztheoretischen, sondern in einem weiten, pluralistischen Sinne verstanden wird. Aber auch hier rennen sie insofern offene Türen ein, als auch ich glaube, dass die Korrespondenztheorie schwerwiegenden Einwänden ausgesetzt ist.[1] Andererseits bin ich sehr viel misstrauischer als sie bezüglich der Akzeptabilität der von ihnen aufgelisteten Alternativen. Das gilt insbesondere für pragmatistische Nützlichkeitstheorien und ‚konstruktivistische' Konzeptionen, die Foucaults Slogan ‚Hervorbringung statt Apophantik' huldigen – mehr dazu unten. Das Wahrheitsverständnis, das ich favorisiere, und das meines Erachtens auch dem Jasperschen Slogan zugrunde liegt, hat zwei Merkmale: Objektivität und Robustheit. Wahrheiten sind ‚objektiv', insofern dieselbe Wahrheit von verschiedenen Menschen gedacht und als wahr anerkannt werden kann; sie sind ‚robust' in dem Sinne, dass sie – im Unterschied zu unserem Fürwahrhalten zu einer bestimmten Zeit – aller menschlichen Macht entzogen sind. Fürwahrhalten ist manipulierbar, Wahrsein ist es nicht. Die Objektivität und Robustheit von Wahrheit scheint mir die Konsequenz eines Prinzips zu sein, das in der Literatur manchmal als ‚Einsicht des Aristoteles' bezeichnet wird. Es besagt, dass die Welt der Dinge und Eigenschaften bei der Erklärung von Wahrheit gegenüber unseren diskursiven Praktiken den Primat hat. Unsere Urteile sind nicht aufgrund menschlicher Praktiken wahr. Erst das So-Sein der Dinge macht sie dazu.[2] Dieser einseitigen explanatorischen Abhängigkeit verdanken Wahrheiten ihre

[1] ‚Korrespondenztheorie der Wahrheit' ist ein Familienname, der verschiedene Wahrheitskonzeptionen umfasst. In seinem enzyklopädischem Standardwerk *Conceptions of Truth* (2003) widmet Künne ihr nur eines von 7 Kapiteln.

[2] ‚Aristotle's insight' – Die Bezeichnung geht auf den amerikanischen Philosophen Roderick Chisholm zurück. Vgl. die Nachweise und die Diskussion in Künne 2003, S, 151 f. – Es ist wichtig, sich von vornherein

Objektivität und ihre „irritierende Hartnäckigkeit“ gegenüber allen Versuchen, sie zum Verschwinden zu bringen.[3]

Wir können die aristotelische Einsicht für unsere Zwecke so formulieren:

Ist das Urteil[4] wahr, dass ein Gegenstand so-und-so ist, dann ist es wahr, *weil* ihr Gegenstand so-und-so *ist*. Das Umgekehrte gilt nicht: Der beurteilte Gegenstand ist nicht so-und-so, *weil* das Urteil, dass er so-und-so ist, *wahr* ist.[5]

Die Hauptpointe dieses Prinzips ist der explanatorische Primat des So-seins des beurteilten Gegenstands gegenüber der Eigenschaft des Urteils, wahr zu sein. Das So-sein des Gegenstands erklärt das Wahrsein des Urteils, nicht das Wahrsein des Urteils sein So-sein. Daraus folgt ein wichtiges Korollar für die Urteilstheorie: Ein wahres Urteil besteht nicht in der (aktiven) *Zuerkennung* einer Eigenschaft, die sein Gegenstand noch nicht hatte, sondern in der (passiven) *Anerkennung* einer Eigenschaft, die er schon hat. Das urteilende Erkennen ist weder ein Erzeugen von Wahrheiten noch der Gegenstände, von denen es handelt.

Auf den ersten Blick mag die aristotelische Einsicht trivial erscheinen. Wie sollte man die Wahrheit eines Urteils sonst erklären, wenn nicht unter Verweis auf die Dinge, von denen es handelt? Ihr kritischer Biss zeigt sich bei Wahrheits- und Urteilstheorien, die mit ihr unverträglich zu sein scheinen. Diese Spannungen möchte ich im Folgenden anhand von drei Wahrheitskonzeptionen skizzieren, die Reinmann und Vohle ausdrücklich als attraktive Alternativen zur Korrespondenztheorie nennen. Die erste ist die pragmatistische Wahrheitstheorie von William James. Die beiden anderen sind die ‚konstruktivistischen‘ Wahrheitskonzeptionen von Bruno Latour bzw. Michel Foucault. Ich hoffe, auf diese Weise den eigentlichen Dissens zwischen den Kommentatoren und mir schärfer herauszuarbeiten und den Dialog voranzubringen.

Sind wahre Urteile wahr, weil sie nützlich sind? – James schreibt in *Pragmatism*: „*Das Wahre ist der Name für alles, was sich als gut erweist, wenn man es glaubt* (1907, S. 59; Hervorh. i. Orig.); „Man kann dann von ihm [i. e. einem wahren Glauben, engl. *belief*] entweder sagen: ‚er ist nützlich, weil er wahr ist‘, oder: ‚er ist wahr, weil er nützlich ist‘. Die beiden Sätze bedeuten exakt dasselbe“ (ibid. S. 135). Lassen wir dahingestellt, ob diese beiden Sätze wirklich „exakt dasselbe“ bedeuten. Der Grundgedanke von James’ konsequentialistischer Wahrheitstheorie ist, dass ein Glaube genau dann wahr ist, wenn er nützlich ist; und er ist nützlich, wenn dass das Festhalten an ihm gute Konsequenzen zeitigt. Demnach ist die Idee eines wahren Glaubens, der zu nichts gut ist, in sich widersprüchlich.

---

klarzumachen, dass die aristotelische Einsicht keine Wahrheitstheorie und erst recht keine Korrespondenztheorie ist.

[3] Arendt 1967, sec. III: Arendt spricht von der „infuriating stubborness“ von Wahrheiten. Das entspricht dem, was ich ihre ‚Robustheit‘ nenne, d. h. ihre Widerständigkeit gegen Manipulation und Macht.

[4] Ich folge hier dem Sprachgebrauch der Logiker seit Aristoteles, die ‚Urteile‘ als primäre Wahrheitswertträger betrachtet haben. Genau genommen ist es nicht der mentale Akt des Urteilens, sondern sein Inhalt, der wahr oder falsch ist.

[5] Ich verallgemeinere hier ein Beispiel von Aristoteles: „Nicht darum nämlich, weil unser Urteil, du seist blass, wahr ist, bist du blass, sondern darum, weil du blass bist, sagen wir die Wahrheit, indem wir dies behaupten“ (*Metaphysik* IX, Kap. 10, 1051b 6-9). Vgl. die Diskussion bei Künne 2003, S. 150 ff.

Eine solche Theorie hat es schwer, die unendlich vielen Wahrheiten zu erklären, die für das menschliche Leben völlig belanglos sind und bleiben. Aber auch diese Schwierigkeit können wir hier ignorieren. Mit Blick auf die aristotelische Einsicht interessiert hier nur James' These, dass der Wahrmacher eines Glaubens seine zukünftige Nützlichkeit sei, d. h. dass ein Glaube wahr ist, *weil* er in Zukunft nützlich zu sein verspricht. Damit stellt James sich gegen Aristoteles, der darauf bestehen würde, dass nicht die Aussicht auf zukünftige Nützlichkeit, sondern das So-sein der Dinge hier und jetzt seine Wahrheit erklärt. Ich begnüge mich an dieser Stelle damit, diese Spannung zu konstatieren.

Erschaffen wahre Urteile ihre Gegenstände? – Mein zweites Beispiel für eine Auffassung, die mit der aristotelischen Einsicht unverträglich erscheint, ist die konstruktivistische Wahrheits- und Urteilstheorie des französischen Wissenssoziologen Bruno Latour. Latour publizierte 1998 einen Essay mit dem Titel ‚Ramsès II est-il mort de la tuberculose?'. Anlass für die Veröffentlichung war eine angebliche Diagnose französischer Pathologen zur Todesursache des ägyptischen Pharaos Ramses II von 1976. Nach eingehenden Untersuchungen seiner Mumie waren die Wissenschaftler – so jedenfalls Latour – zu dem Urteil gelangt, dass der Pharao vor über 3000 Jahren an Tuberkulose starb.[6] Latour widerspricht diesem Befund – allerdings nicht mit einem medizinischen Gegenargument, das eine andere Todesursache favorisiert. Vielmehr hält er den Medizinern entgegen, dass ihre Diagnose schon deshalb irrig sein müsse, weil der Tuberkulosebazillus erst mehr als 3000 Jahre nach dem Tod von Ramses II in einem Berliner Labor von Robert Koch *erschaffen* wurde. Der 24. März 1882 wird jedes Jahr als Tag der *Entdeckung* des TB-Erregers gefeiert. Aber nach Latour ist *mycobacterium tubercolosis*, wie der Erreger heute offiziell genannt wird, Kochs Kopfgeburt. Koch habe den Einzeller durch sein Urteil buchstäblich in die Welt gebracht: „Vor Koch hatte der Bazillus keine reale Existenz" (ibid., S. 84). Da *mycobacterium tubercolosis* in Ramses' Todesjahr 1213 v. u. Z. noch nicht existierte, so Latours Folgerung, konnte er nicht an TB gestorben sein. Der Befund der Pathologen sei so anachronistisch wie der, dass der Pharao „von einer Maschinengewehrsalve" niedergestreckt worden sei (ibid., S. 84). Latours Resümee verallgemeinert diesen Punkt für wissenschaftlichen Erkenntnisse überhaupt. Wissenschaftler:innen, so Latour, „begnügen sich nicht mit dem Ent-decken: Sie produzieren, sie stellen her, sie bauen."[7] Hätte Latour recht, dann wäre die Verleihung des Nobelpreises an Koch meines Erachtens ein moralischer Skandal ersten Ranges. Stattdessen hätte man ihn wegen Verbrechen gegen die Menschlichkeit vor

[6] In Wahrheit haben die Pariser Pathologen 1976 keine Diagnose zur Todesursache von Ramses II gestellt. Der über 90jährige Pharaoh starb wahrscheinlich an Altersschwäche in Verbindung mit einer Sepsis. Der Sache nach irrt Latour, aber wir können seine Frage so verstehen: Ist es möglich, dass der Pharaoh an TB starb?

[7] „… les chercheurs ne se contentent pas de ‚dé-couvrir': ils produisent, ils fabriquent, ils construisent" (Latour 1998, S. 84). Der Bindestrich in 'dé-couvrir' ist wahrscheinlich eine Anspielung auf Heideggers ‚entdecken. – Bernard Williams wundert sich zurecht über Latours „bemerkenswerte Annahme, die Befähigung der Wissenssoziologie zur Aufstellung wahrer Behauptungen über die Naturwissenschaften sei größer als die Befähigung der Naturwissenschaften zur Aufstellung wahrer Behauptungen über die Welt" (Williams 2013, S. 13).

Gericht stellen sollen.[8] Auch hier belasse ich es bei der Feststellung, dass Latours konstruktivistische Theorie mit der aristotelischen Einsicht unverträglich scheint.

Erschaffen wahre Urteile Wahrheiten? – Foucault stellt die aristotelische Einsicht auf den Kopf, wenn er sagt: „Die Wahrheit gehört … nicht zur Ordnung dessen, was ist, sondern dessen, was eintritt: Ereignis. Sie wird nicht festgestellt, sondern hervorgerufen: Hervorbringung statt Apophantik“ (Foucault 2014, S. 854). *Contra* Aristoteles sei ein wahres Urteil kein *logos apophantikos*, d. h. ein mentaler Akt, der feststellt, was unabhängig vom Urteil der Fall ist, weil es der Fall ist. Vielmehr sei ein wahres Urteil ein Ereignis, dass seine eigene Wahrheit hervorbringt. Auch in Foucaults Augen, so scheint es, hat Koch durch sein Urteil vom 24. März 1882 etwas Neues geschaffen. Vielleicht nicht *mycobacterium tubercolosis* selbst, aber die *Wahrheit*, dass er der TB-Erreger ist. Diese Wahrheit ist Kochs Kopfgeburt und ihr Geburtstag ist der 24. März 1882.[9]

Ist Foucaults Wahrheitserzeugung plausibler als Latours Gegenstandserzeugung? Betrachten wir noch einmal die Frage nach der Todesursache von Ramses II. Wäre Foucaults These korrekt, dann war Kochs Wahrheit zum Zeitpunkt der Untersuchung der Pariser Forensiker im Jahr 1976 gerade mal 96 Jahre alt, während Ramses II schon über 3000 Jahre tot war. Gleichwohl zögerten die Wissenschaftler nicht, dieser relativ jungen Wahrheit eine Schlüsselrolle in der Begründung ihrer Diagnose einzuräumen. Sie argumentierten in etwa so:

(1) Im Gewebe der Mumie finden sich genetische Reste von *mycobacterium tubercolosis.*

(2) *Mycobacterium tubercolosis* ist der Erreger von TB.

Also: (3) Es ist wahrscheinlich, dass Ramses II an TB litt und daran starb.

Auf den ersten Blick wirkt das Argument überzeugend. Die Wahrheit von (1) und die Wahrheit von (2) machen die Wahrheit der Konklusion (3) zumindest wahrscheinlich. Aber aus Foucaults Sicht weist das Räsonnement der Pathologen eine offenkundige Schwachstelle auf. Denn Prämisse (2) war im Todesjahr von Ramses II (noch) nicht wahr. Es war nicht der Fall, dass *mycobacterium tubercolosis* vor 3000 Jahren der Erreger von TB war. Aber dann kann der Pharao nicht an TB erkrankt und daran gestorben sein. Auch die konstruktivistische Wahrheitsauffassung Foucaults scheint also das Ergebnis von Latour der Sache nach zu bestätigen.

Die obigen Überlegungen sollen nur zeigen, dass die aristotelische Einsicht, wenn sie denn eine ist, mit den Wahrheits- und Urteilstheorien von James, Latour und Foucault unverträglich zu sein scheint. Ob diese Unverträglichkeit tatsächlich besteht oder ob sie mit der aristotelischen Einsicht versöhnt werden können, muss hier offenbleiben. Sollten sie sich jedoch als unvermeidlich erweisen, dann müssten wir zwischen ihnen und dem aristotelischen Prinzip wählen. Angesichts der kontraintuitiven Konsequenzen der Theorien von

[8] Kaum hatte ich diese Zeilen geschrieben, musste ich erstaunt feststellen, dass Künne in einem jüngst erschienenen Aufsatz in ähnlichen Worten denselben Punkt macht: Künne 2024, S. 311.

[9] Künne weist darauf hin, dass Heidegger 1927 eine ähnliche Auffassung vertrat: „Bevor die Gesetze Newtons entdeckt wurden, waren sie nicht wahr; daraus folgt nicht, dass sie falsch waren … Die Gesetze Newtons waren vor ihm weder wahr noch falsch … [Sie] wurden durch Newton wahr“ (Heidegger 1927, S. 309).

James, Latour und Foucault fällt mir die Entscheidung leicht. Aber ich vermute, dass der eigentliche Dissens zwischen mir und den Kommentatoren an dieser Stelle manifest wird. Reinmann und Vohle scheinen mir die aristotelische Einsicht mit dem Badewasser der Korrespondenztheorie ausschütten zu wollen. Ich hingegen glaube, dass sie tatsächlich eine Einsicht ist und den *common sense* auf ihrer Seite hat. Da es bei Wahrheitstheorien darum geht, unser aller implizites Verständnis von ‚wahr' explizit zu machen, scheint mir das zu bestätigen, dass Aristotles hier den Nagel auf den Kopf trifft. Der von ihm behauptete explanatorische Primat der Welt bildet meines Erachtens auch die Grundlage für die Objektivität und Robustheit von Wahrheit. Aber dieser Nachweis würde den engen Rahmen einer Replik sprengen.

Literatur

Aristoteles: *Metaphysik*, Hg. H. Seidel, Felix Meiner Verlag 1989.

Arendt, H.: ‚Truth and Politics' *The New Yorker*, 25. Februar 1967.

Foucault, Michel: *Schriften in vier Bänden. Dits et Ecrits*. Band II, 1970-75, Hg. Daniel Defert & al., Suhrkamp 2014.

Heidegger, M.: *Sein und Zeit*, Halle an der Saale 1927.

James, W. *Pragmatism: A New Name for Some Old Ways of Thinking*. Longmans, Green, and Co., 1907.

Künne, W.: *Conceptions of Truth*, Oxford University Press 2003.

--"--: ‚Veritas filia temporis? Reflexionen über Wahrheit und Zeit, angeregt von einem Teppich' in: M. Hoffmann und T. Martin, Hg. *Ebenen des Verstehens. Wolfgang Künne im Gespräch*, Brill 2024.

Latour, B.: ‚Ramsès II est-il mort de la tuberculose?' La Recherche, Nr. 307, 1998, S. 84–85.

Williams, B.: *Wahrheit und Wahrhaftigkeit*, Suhrkamp 2013.

Die Zukünfte der Universität: zwischen technisierter und lebendiger Wissenschaft

Dominikus Herzberg

In diesem Text gehe ich davon aus, dass die Idee der Universität bzw. allgemein der Hochschulen etwas mit Wissenschaft zu tun hat – Stichwort „Bildung durch Wissenschaft". Es bedarf einer Klärung des Verhältnisses von KI, Sprache und Wissenschaft, weil Textproduktionsmaschinen wissenschaftliche Form erzeugen können, ohne wissenschaftlichen Weltaufschluss zu vollziehen. Wo diese Differenz übersehen wird, droht Scheinwissenschaft – und mit ihr eine Gefährdung der Idee einer Bildung durch Wissenschaft.[1]

Künstliche Intelligenz ist nicht das Problem

Für eine angemessene Einordnung, Diagnose und Zukunftsprojektion zunächst eine Klärung: Wenn von generativer oder, allgemein, von Künstlicher Intelligenz (KI) gesprochen wird, so hilft das nicht bei der Einkreisung des Problems, das sich derzeit an den Hochschulen bemerkbar macht. KI gibt es schon lange, die Anfänge lassen sich bis in die 1950er Jahre hinein verfolgen. KI als eine besondere Klasse algorithmisch-rechnerischer Verfahren wird in vielen Wissenschaften seit langem eingesetzt und genutzt – und das hat nicht die geringste Krise in Lehre, Forschung und Entwicklung ausgelöst. Dass Rechenmaschinen im Allgemeinen Leistungen erbringen, die mechanisch-maschinell faktisch nicht umsetzbar sind, sondern der Digitaltechnik als Grundlage bedürfen, setzt Datafizierung samt Rechenvollzügen zu Sinnlichkeit und kognitiven Vollzügen in gewisser Hinsicht in Konkurrenz – und dann doch wieder nicht. Das Konkurrenzverhältnis wird mit dem Begriff der künstlichen Intelligenz markiert, mehr nicht. Dass ein Computer sehr viel besser Schach spielen kann als viele (menschliche) Spitzenspielerinnen und Spitzenspieler, das kann man als technologischen Erfolg bezeichnen, nimmt aber nicht unbedingt den Schachspielenden die Freude und den Wettbewerb, sich in dem Spiel zu messen. Wohl aber haben sich die Vorbereitung und das Training im Schachspiel verändert: mit der Schachmaschine geht das intensiver und systematischer, in ihr findet sich ein unermüdlicher und unerschöpflicher Partner, der es erlaubt, das Training an investierter Zeit und an Intensität auf die Spitze zu treiben.

KI ist als Begriff zu grob. KI ist immer gegenstandsbezogen, sie bezieht sich auf einen speziellen Bereich operativ vollziehbarer Verhältnisse. Man muss das Augenmerk darauf lenken, *was* für eine operative Leistung maschinell mit Verfahren der KI durch eine (vernetzte) Rechenmaschine erbracht wird. Beim Schachcomputer oder einem Schachprogramm betrifft dies das operative Feld des Schachspielens, bei ChatGPT und Co. hingegen die anschlussfähige Textproduktion bzw. Textgeneration. Um dieses operative Feld

[1] Ich habe mich sehr darüber gefreut, dass Markus Stepanians diesen Textbeitrag zum Symposium „Zur Idee der Universität im 21. Jahrhundert: Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz" für die Finalfassung mit Kommentierungen und einigen Formulierungsvorschlägen bedacht hat, die ich alle sehr gerne übernommen habe. Vor allem spreche ich nun von Textproduktionsmaschinen statt von Sprachmaschinen.

begrifflich zu fassen, spreche ich von Textproduktionsmaschinen, auch wenn der Ausdruck etwas holprig klingen und ein wenig sperrig zu lesen sein mag.

Was derzeit das Hochschulsystem erschüttert oder zumindest erheblich irritiert, ist nicht KI im Allgemeinen, sondern es sind die Textproduktionsmaschinen im Besonderen. Es geht um das Eindringen der Maschine in ein operatives Feld, das bislang allein Domäne des Menschen gewesen ist: die Sprache. Die Fortschritte der letzten zwei Jahrzehnte in der Spracherkennung und der Sprachsynthese erlauben zudem den Austausch über das gesprochene Wort. Der wesentliche Punkt ist jedoch ein anderer: Textproduktionsmaschinen berücksichtigen bei der Textproduktion den vorangegangenen Textkontext, und zwar tun sie das in bemerkenswert anschlussfähiger Weise, nicht nur inhaltlich wie auch stilistisch, sondern dem Anschein nach sogar semantisch. Die z. B. durch die Sprache vermittelte Anfrage oder Aufforderung wird als solche erkannt und bearbeitet, die Frage wird scheinbar beantwortet und der Aufforderung wird sprachlich nachgekommen. Das geht so weit, dass Aufforderungen, wie etwa eine bestimmte Textgattung zu produzieren oder Erklärungen zu einem Sachverhalt zu liefern, erfüllt werden. Derartig hervorgebrachte Texte und Dialoge könnten auch die von Menschen sein. Das charakterisiert das Spektrum, wie weitreichend das operative Feld der Sprache von Textproduktionsmaschinen beherrscht wird. Da Textproduktionsmaschinen an unvorstellbar großen Textmengen trainiert werden, haben sie zu allem was zu sagen und können in vielen Bereichen scheinbar Auskunft geben – nicht durch Simulation kognitiver Prozesse, sondern allein durch die erlernten strukturellen Verhältnisse in Sprache. Der „Strukturalismus“ geht auf den Schweizer Linguisten Ferdinand de Saussure zurück. Sein Grundgedanke ist, dass die Bedeutung sprachlicher Zeichen allein durch die *rein syntaktischen* Beziehungen bzw. Differenzen der sprachlichen Zeichen untereinander und ihr Gebrauchsverhältnis hervorgeht, nicht jedoch durch ihren außersprachlichen Weltbezug, d. h. ihre semantische Referenz auf äußere Dinge. Die sogenannte „Transformer-Architektur“ für künstliche neuronale Netze, welche die assoziativen und positionalen Beziehungen von Wortbruchstücken bis hin zu Verwendungsarten und -mustern von Sprache nachbildet, kann als Umsetzung dieses strukturalistischen Grundgedankens angesehen werden (Weatherby, 2025).

Textproduktionsmaschinen entgehen der Verwicklung der Sprache mit der Welt

Wie wir mittlerweile wissen – und das ist Teil des Erfolgs –, ist diese generative Texterzeugung bemerkenswert ergiebig und nützlich. Auf den ersten Blick sind Textproduktionsmaschinen mit ihren rein syntaktischen Strukturmodellen eine empirische Bestätigung der Ideen von Saussure. Es ist und bleibt überraschend, wie umfassend Sprachmodelle mit Sprache umgehen können. Dem Anschein nach kann man sie genau wie menschliche Quellen dazu nutzen, um Wissenslücken zu schließen, Weltbeschreibungen zu produzieren, Lebensprobleme zu diskutieren und vieles mehr. Das scheint eine hocheffektive, technisierte Ausbeutung von Sprache als zugrunde liegendem Trainingsmaterial zu sein.

Auf den zweiten Blick muss es jedoch rätselhaft erscheinen, wie rein syntaktisch operierende Textgeneratoren ohne Weltbezug zu diesen Leistungen in der Lage sein sollen. Sind wir berechtigt, den rein syntaktisch-strukturellen Output einer KI als wirkliche *Antworten* auf unsere Welt- und Lebensfragen zu verstehen? Für *menschliches* Denken, Handeln und

Sprechen ist der Weltbezug wesentlich. Gemäß Wilhelm von Humboldt ist Sprache konstitutiv zur Ausbildung eines Selbst- und Weltverhältnisses, und Heidegger geht darüber noch einmal hinaus: Sprache ist elementar im Daseinsverhältnis mit der Welt (vgl. Posselt und Flatscher, 2018). Menschliche Fragen sind Fragen über die Welt und das Leben in ihr, während den „Antworten" der KI der Weltbezug abgeht. Verstehen wir ihre Antworten kontrafaktisch dennoch so, als ob sie von denselben Dingen handelten wie unsere Fragen, dann schreiben wir ihnen einen Sinn zu, den rein syntaktische Strukturen von sich aus nicht haben. Diese Sinngebungen erfolgen auf eigene Verantwortung und eigenes Risiko.

Das Interessante ist, dass der Strukturalismus in seiner Reinform Grenzen der Beschreibbarkeit und der Erklärbarkeit mit sich bringt. Einer Textproduktionsmaschine liegt zwar ein Sprachmodell zugrunde, doch ist das Modell gleichzeitig seine Erklärung: Es repräsentiert in seiner Komplexität als Konfiguration eines neuronalen Netzwerks ein kaum tiefer ergründbares Strukturabbild. So spannend die Erkenntnisse der sogenannten Interpretabilitätsforschung zu Sprachmodellen sind, man ist weit davon entfernt zu verstehen, wie genau Textproduktionsmaschinen funktionieren und was ihre inhärenten Mechanismen und damit Möglichkeiten und Beschränkungen sind. Es ist hingegen dennoch möglich, sie nutzbar zu machen und ihre Nützlichkeit zu steigern. Das ist ein typisches Merkmal für technikwissenschaftliche Zugänge: Das Zielkriterium ist Effektivität, die Wertschöpfung ist zentral. Die ingenieursmäßige Konstruktion, Komposition und Nutzbarmachung stehen im Vordergrund – dafür muss nicht alles verstanden werden. Textproduktionsmaschinen müssen nur funktionieren, das reicht.

Struktur und Bedeutung in Sprache und in den Wissenschaften

Der Strukturalismus ist nicht allein auf Sprache beschränkt, sondern kann als generelles Instrument einer gewissen Form der Gewinnung von Erkenntnis verstanden werden. So wie es die Entwicklung der Textproduktionsmaschinen vorgemacht hat, kann man den wissenschaftlich zu untersuchenden oder zu beforschenden Gegenstand in seinen Strukturen versuchen, durch rechnerische Vollzüge mit Techniken der künstlichen Intelligenz zu erfassen und für generative, prädiktive, simulative etc. Anwendungsfälle nutzbar zu machen. Voraussetzung dafür ist das Vorhandensein von ausreichend großen Datenmengen zum Forschungsgegenstand. Dies ist der oben angemerkte Gebrauch von KI, der seit jeher unproblematisch ist.

Die Erkenntnisgewinnung kann aber auch eine ganz andere Seite haben, nämlich, wenn zu wenig oder deutlich zu viele Daten vorliegen, wenn Strukturverhältnisse einer Domäne verstanden werden wollen, wenn Strukturen in Modelle einfließen sollen, die durch Abstraktion und Reduktion komplizierte Zusammenhänge auf wesentliche Anteile beschränken, wenn Hypothesen aufgestellt, wenn Annahmen und Ideen zur Erklärung, zur Erweiterung oder zum Transfer aufgebracht werden sollen, wenn also Theoriebildung stattfindet. Diese Art der Erkenntnis, ich nenne sie hier theoriebildende Erkenntnis, steht der strukturellen Erkenntnis gegenüber.

Die nachstehende Tabelle stellt die Sprache im Speziellen und die Wissenschaft im Allgemeinen in ihren Polen, die ich zugleich als Konkurrenzverhältnisse deuten möchte, von Struktur durch Rechenvollzüge und von Weltbezug durch Handlungsvollzüge dar. Der Handlungsbegriff ist dabei so weit zu verstehen, dass er auch Sprechakte einschließt.

	Struktur durch Rechenvollzüge	Weltbezug durch Handlungsvollzüge
Sprache	Textproduktionsmaschine	Weltaufschluss
Wissenschaft	KI-Modell	Theoriebildung

Die Tabelle führt nicht absichtslos Sprache und Wissenschaft auf. Beide stehen in einem besonderen Verhältnis zueinander: Was die Sprache angeht, so ist Wissenschaft die Unternehmung, alles in Sprache (mit)teilbar und in Sprache (be)denkbar zu machen. Wissenschaft besteht nicht nur im privaten Erkenntnisgewinn durch Urteile, sondern in der sprachlichen Mitteilung dieser Urteile in Behauptungen. Aber Sprache ist kein bloßes Mittel der Kommunikation von Gedanken, sondern macht Denken im eigentlichen Sinne erst möglich.

Kurioserweise ist es durch dieses intrinsische Verhältnis von Wissenschaft und Sprache überhaupt erst möglich geworden, dass Textproduktionsmaschinen wissenschaftlich anschlussfähige Texte hervorbringen können. Der Anspruch, ein operatives Feld der Wissenschaft sprachlich (be)denkbar zu machen, erfordert klare Begrifflichkeiten und Vorstellungsbilder. Wenn möglich, werden Formalisierungen angestrebt und genutzt, und es wird auf Nachvollziehbarkeit geachtet. All das schlägt sich in syntaktischen Formen nieder, die wieder und wieder in jeder Publikation vollführt werden und die einer Textproduktionsmaschine hinreichend erlernbare "Strukturen" anbieten, um damit anschlussfähig auch in wissenschaftlichen Zusammenhängen Texte hervorbringen zu können.

Zwischen technisierter und lebendiger Wissenschaft

Die obige Tabelle erlaubt einen Ausblick auf zwei radikale Wissenschaftsformen, und zwar auf eine hochgradig technisierte und formalisierte Wissenschaft (erste Spalte) und eine, die ich als lebendige, auf die Lebenswelt bezogene Wissenschaft bezeichne (zweite Spalte). Zwischen diesen beiden Formen gibt es selbstverständlich Varianten mit unterschiedlichen Formanteilen.

Dadurch, dass Sprache durch Textproduktionsmaschinen als technisches Mittel zur Verfügung steht und automatisiert genutzt werden kann, lassen sich nun sprachlich mediatisierte Schnittstellen in Prozesse der Forschung und Entwicklung und der Erkenntnisgenerierung einführen. Damit ist der ingenieursmäßige Zugriff auf Sprache in Kombination mit technischen Apparaten möglich. Das heißt, kurz gesagt: Textproduktionsmaschinen können als Agenten auch in Forschungs-, Entwicklungs- und Erkenntnisprozessen wirken und sind dazu in ingenieursmäßiger Manier zu konfigurieren und zu orchestrieren. Die Informatik macht das im Moment vor und versucht sich am agentisch gesteuerten Entwickeln von Software (vgl. Kim und Yegge, 2025). Da Textproduktionsmaschinen aus Ingenieurssicht eine Risikotechnologie sind, muss das Risiko gemanagt und durch Qualitätssicherungsprozesse und -mittel eingedämmt und beherrschbar gemacht werden. Das ist – wie sich derzeit zeigt – eine nicht triviale Herausforderung.

Eine Wissenschaft, die sich anhand von (sprachlichen und nicht-sprachlichen) Handlungsvollzügen Welt aufschließt und Theoriebildung betreibt, zeigt eine andere Form. In diesem

Sinne ist Wissenschaft eine Unternehmung, die in der Eigenlogik ihrer Disziplin bzw. ihres operativen Feldes Handlungsvollzüge durchführt, die mit dem Gegenstand des Interesses bzw. der Untersuchung einen Umgang finden, und immer davon ausgeht bzw. davon ausgehen muss, dass dieser Gegenstand im Handeln widerständig ist, die bisherigen Beschreibungslogiken unzureichend sind und den Erkenntnisgegenstand nicht hinreichend fassen. Diese Widerständigkeit und dieses Ringen mit dem Gegenstand des Interesses bezeichne ich als lebendige Wissenschaft.

Diese beiden Formen sind für sich genommen unproblematisch als Erscheinungsformen des Wissenschaftlichen. Neu wird in Zukunft die radikal technisierte und hochgradig automatisierte Durchführung von Forschung und Entwicklung sein.

Das Problem entsteht dadurch, dass sich die eine Wissenschaftsform bei der anderen Erkenntnisform bedient oder sie ersetzt und ihr originäres Beziehungsverhältnis zu Sprache durcheinander bringt: Wenn die Textproduktionsmaschine als Weltaufschluss und/oder Theoriegenerator missverstanden wird, wenn der Weltaufschluss durch eine Textproduktionsmaschine und/oder ein KI-Modell erschlossen wird, wenn ein KI-Modell mit Theoriebildung und/oder Weltaufschluss verwechselt wird, wenn Theoriebildung durch ein KI-Modell und/oder eine Textproduktionsmaschine geleistet wird – dann korrumpieren sich diese beiden Wissenschaftsformen.

	Struktur durch Rechenvollzüge	Weltbezug durch Handlungsvollzüge
Sprache	Textproduktionsmaschine als Weltaufschluss und Theoriegenerator	Weltaufschluss durch Text- produktionsmaschine und KI-Modell
Wissenschaft	KI-Modell als Theoriebildung und Weltaufschluss	Theoriebildung durch KI-Modell und Textproduk- tionsmaschine

Dann funktionieren diese beiden Wissenschaftsformen nicht mehr, sondern produzieren Scheinwissenschaft, die lediglich der sprachlichen Form nach wissenschaftlich erscheinen mag. Und sie korrumpieren ihr Verhältnis zur Sprache, indem Bedeutungszuschreibungen durch Textproduktionsmaschinen nicht hergestellt, sondern nur generiert werden, und die Produktionen von Textproduktionsmaschinen für das Ergebnis „intelligenter“ Maschinen mit Weltbezug gehalten werden. In dem Moment wäre die Programmatik einer „Bildung durch Wissenschaft“ gescheitert.

Die Aufgabe der Universität bestünde dann nicht darin, Textproduktionsmaschinen entweder zu integrieren oder fernzuhalten. Ihre Aufgabe bestünde darin, die Differenz zwischen struktureller Verfügbarkeit und lebendigem Weltaufschluss lehrbar, erfahrbar und kritisierbar zu machen. Genau darin könnte eine zeitgemäße Wissenschaftsdidaktik bestehen: nicht in der bloßen Vermittlung von KI-Kompetenz, sondern in der Unterscheidungsfähigkeit, wann Sprache wissenschaftliche Erkenntnis trägt und benötigt – und wann sie nur deren Form simuliert.

Literatur

Kim, G. & Yegge, S. (2025). *Vibe Coding: Building Production-Grade Software With GenAI, Chat, Agents, and Beyond*. IT Revolution.

Posselt, G. & Flatscher, M. (2018). *Sprachphilosophie: Eine Einführung*. 2. Auflage. facultas.

Weatherby, L. (2025). *Language Machines: Cultural AI and the End of Remainder Humanism*. University of Minnesota Press.
https://www.upress.umn.edu/9781517919320/language-machines/

Ach du liebe Wahrheit – eine Nachbetrachtung aus dem Dialog mit Markus Stepanians

Dominikus Herzberg

Wahrheit ist ein schönes Wort, ich mag es. Markus Stepanians, er ist Philosoph, brachte die Wahrheit für das Hamburger Symposium „Zur Idee der Universität im 21. Jahrhundert: Humane Hochschulbildung in Zeiten von Künstlicher Intelligenz“ ins Spiel. Sie solle nach wie vor der zentrale Wert der Universität sein, so meinte er, in diesem Jahrhundert und, so habe ich ihn verstanden, – zeitlos im Anspruch – eigentlich für immer.

Bildung ohne Wahrheit ist schwer vorstellbar. Jedoch begegnet mir die Wahrheit im akademischen Alltag eher in ihrem Gegenteil, der Unwahrheit. Die Unwahrheit, wie auch die Unwahrhaftigkeit, spielen im Prüfungsrecht, beim akademischen Fehlverhalten und in der Konfrontation von Wissenschaft und Gesellschaft eine Rolle. Wenn es zum Beispiel um Täuschung, Betrug, Plagiat, Fälschung, Irreführung, Desinformation und die Lüge geht. Über die Unwahrheit lässt sich viel leichter reden. Aber wie ist es nun mit der Wahrheit?

Als Mitglied einer Hochschule für Angewandte Wissenschaften (HAW) läuft mir die Wahrheit selten über den Weg. Das liegt nicht so sehr an der Hochschulform, sondern an der bei ihr vorrangig ausgeübten Wissenschaftsform: der angewandten Wissenschaft. Die Technikphilosophie, nehmen wir sie als prototypischen Zugang zu den angewandten Wissenschaften, basiert auf einem ganz anderen Erkenntnisbegriff, nämlich dem der Effektivität: Es geht darum, dass etwas funktioniert. Ausgangspunkt ist der pragmatische Syllogismus, der eine elementare handlungstheoretische Denkfigur darstellt (Kornwachs, 2012, 2013). Praktisches Wissen entsteht aus der Beobachtung, dass ein Effekt E mit einer Handlung H erreicht werden kann. Ist das mit einer gewissen Zuverlässigkeit der Fall, liegt ein Wissen um den Zusammenhang E per H vor; man nennt dieses Wissen auch eine Regel. Eine solche Regel muss nicht wahr sein, sie muss sich lediglich als effektiv erweisen. Technologisches Wissen entsteht, wenn ein Zweck oder eine Norm mit E als Ziel über eine Zweck-Mittel-Relation verknüpft ist:

Wenn H zu E führt, notiert als Implikation $H \rightarrow E$, und wenn E das Ziel ist, dann tue oder versuche E per H. Dabei gibt es eine subtile Unterscheidung: Der Effekt und die Handlung als etwas zu Beobachtendes sind in Fettschrift notiert (E per H), ihre formale Abbildung in ein Zeichensystem als $H \rightarrow E$ ohne Fettschreibung. So werden die Bereiche des Handelns in der Welt und des formalen Umgangs mit Symbolen in Bezug gesetzt, aber eben nicht gleichgesetzt (Herzberg, 2023, S. 10f.). Empirie und Theorie kommen hier auf ganz eigene Weise zusammen: Der Test als empirische Überprüfung und die Simulation als formale Berechnung sind erkenntnistheoretisch tragend. Die formal-theoretisch hergeleitete Funktionsvermutung trifft auf die empirische Überprüfung ihrer Effektivität. Es geht dabei nicht um eine Verallgemeinerung, sondern um den wiederholbaren Einzelfall. Effektivität basiert auf einem pragmatischen Wahrheitsbegriff, der Beobachtung und Effekt in Zusammenhang bringt.

Dieser erkenntnistheoretische Zugang ist auch auf meine Disziplin, die Informatik, anwendbar, jedoch mit einem interessanten Unterschied: In abstrakt-operativer Sicht (wenn man die materielle Grundlage eines Computers ignoriert) fallen das formale und das handelnde System zusammen. Ein Computer ist lediglich der aberwitzig viel schneller ausführende formale Mechanismus, den ein Mensch prinzipiell auch per Stift und Papier symbolisch durchführen könnte – ihre Ergebnisse wären gleich. Der Computer vermag programmgesteuert formale Handlungen zu vollziehen. Ob und wie diese Programme sinnhafte Abbilder oder gar gelungene Täuschungen von außerrechnerischen Weltbezügen sind, ob Modelle adäquat dem Gegenstand ihres Anlasses entsprechen, das sind Fragen, auf die die Informatik erkenntnistheoretisch (nach meinem Wissensstand) noch keine Antwort gefunden hat. Zweifelsfrei ist lediglich, dass das Testen von Programmen und Modellen eine essenzielle empirische und formale Nahtstelle ist.

Die generative KI, die Textproduktionsmaschinen, die uns beim Symposium vorrangig beschäftigt haben, sind an schier unvorstellbar großen Textmengen trainiert worden, und sie sprechen bemerkenswert menschengleich zu uns und mit uns. Sie sind technisch hervorgebrachte Erzeugnisse. Ihr Sprachvermögen ist im statistischen Ansatz maschineller Lernverfahren begründet, allerdings bedurfte es eines neuen und besonderen Aufbaus, der sogenannten Transformer-Architektur, darauf weist Weatherby (2025) hin. Das Sprachvermögen kommt darüber hinaus ohne explizite sprachtheoretische Annahmen aus; es basiert allein auf dem verarbeiteten Trainingsmaterial und einer Einhegung der Maschine im Umgang mit Menschen durch ein Finetuning. Im Grunde analysieren diese Maschinen nur die kleinen und großen Strukturen, wie Wörter und Satzzeichen in den Trainingstexten verwendet werden. Auf dieser Basis bauen sie dann neue Texte zusammen. Klug erscheinende Antworten sind darin genauso angelegt wie das Phantasma und der Unsinn. Wer hier nach Intelligenz oder Wahrheit sucht oder beides einzubauen hofft, findet dazu keine Ansatzpunkte. Wahrheit gehört nicht zu den konstitutiven Kriterien, nach denen ein Sprachmodell bei der Textproduktion operiert. Sprache wird hier im Medium der Sprache berechnend vollzogen. Warnke (2026) weist zurecht darauf hin, dass Sprachmodelle trotz ihrer Neuronen-Metapher nichts mit dem Aufbau des menschlichen Gehirns gemein haben. Vielmehr sieht er die Linguistik als Bezugswissenschaft, die sich mit der Sprache und ihren Möglichkeiten und Grenzen von Weltbezügen und mit der Semantik befasst. Es sei auf den Strukturalismus von Saussure verwiesen, so z.B. bei Herzberg (im Druck), und auf den darauf aufsetzenden Distributionalismus des Linguisten Harris, so bei Warnke (2026).

Aus technikwissenschaftlicher Sicht (vgl. Poser, 2012) erfüllen Textproduktionsmaschinen eine Funktion, und sie erfüllen einen Zweck im Kontext eines Ziels. Dazu müssen aber auch die Beziehungen zu Werten hergestellt werden. Nach Poser müssen Technikтheorien "Vier-Ebenen-Theorien sein", es gilt "Kriterien für die ökonomischen, sozialen, psychischen und ökologischen Bedingungen zu formulieren, die Techniken neben ihrer technischen Effektivität und Effizienz zu erfüllen haben" (S. 330). Diese unaufgeregte, jedoch substanzielle Betrachtungsweise vermisse ich, wenn es um den Einsatz von Textproduktionsmaschinen im Hochschulkontext geht. Die ökonomische Ebene tritt oft als Effizienz, als Einsparungspotenzial bei der Produktion und Bewältigung von Textarbeit auf. Die

Effizienz fragt nicht, ob ein Ziel erreicht wird, sondern mit welchem Aufwand. Die Prüfung des Weltbezugs eines generierten Textes, sein gründliches Durchdenken, würde diesen Aufwand zurückholen. Vordergründig mag es effizient sein, das zu unterlassen. Wo jedoch Wahrheit oder Weltbezug zum Zweck der Anwendung einer Textproduktionsmaschine gehören, müssen Kriterien der Zielerreichung von außen an das technische Verfahren herangetragen und empirisch geprüft werden.

Was folgt aus dieser Nachbetrachtung? Geht es also um Wahrheit vs. Effektivität? Nein, diese Zuspitzung geht zu weit. Für den Umgang mit Textproduktionsmaschinen in der Hochschule liefert mir der technikwissenschaftliche Zugang über die Effektivität hilfreiche Anknüpfungspunkte, um daran bildungstheoretische Überlegungen anschließen zu können. Beim Gebrauch von Textproduktionsmaschinen ist die Wahrheit rasch eines ihrer ersten Opfer, denn diese Maschinen sind konstruktionsbedingt keine Wahrheitsapparate. Die Nutzung von Textproduktionsmaschinen in Bildungszusammenhängen muss je nach Zweck und Kontext wichtige Kriterien der Zielerreichung mit einbeziehen; ihre Gewährleistung und Überprüfung können nicht von der KI kommen. Bei einer Zusammenfassung eines Textes durch eine Textproduktionsmaschine sind andere Kriterien anzulegen als beim Programmieren mit einer solchen Maschine oder bei der Assistenz, sich etwas erklären zu lassen. Es muss genau hingeschaut werden, wo z. B. Wahrheit, die eigene Stimme oder die eingebrachte Erfahrung eine Rolle spielt oder spielen soll – und wie wichtige Bildungsanliegen wie Selbstbestimmung, Wissens- und Kompetenzaufbau im Sinne der Vier-Ebenen-Theorie berücksichtigt werden. Wo das unterbleibt, geht es wieder um Täuschung, Betrug, Plagiat, Fälschung, Irreführung, Desinformation und die Lüge. Bildung scheint ohne Wahrheit und Wahrhaftigkeit nicht zu funktionieren, die Textproduktionsmaschine hingegen schon.

Literatur

Herzberg, D. (2023). Konzeption einer Wissenschaftsdidaktik für die Fachhochschule [Dissertation]. Universität Hamburg

Herzberg, D. (im Druck). Die neue Herausforderung in der Hochschullehre: Das Problem ist nicht die KI, sondern die Bedeutung von Sprache. In J. Noller et al. (Hrsg.), *Lehren und Lernen in Mensch-Maschine-Interaktion*, Reihe "Perspektiven der Hochschuldidaktik", Springer VS.

Kornwachs, K. (2012). *Strukturen technologischen Wissens: Analytische Studien zu einer Wissenschaftstheorie der Technik*. edition sigma

Kornwachs, K. (2013). *Philosophie der Technik: Eine Einführung*. C.H. Beck

Poser, H. (2012). *Wissenschaftstheorie: Eine philosophische Einführung*. 2. Aufl. Reclam

Warnke, M. (2026). *Large Language Kabbala: Eine kleine Geschichte der Großen Sprachmodelle*. Matthes & Seitz

Weatherby, L. (2025). *Language Machines: Cultural AI and the End of Remainder Humanism*. University of Minnesota Press

Wahrheitsapparate und Effektivität – zu zwei Thesen von Dominikus Herzberg

Markus Stepanians

Dieser Kommentar ist angeregt durch zwei Thesen von Dominikus Herzberg in ‚Die Zukünfte der Universität – zwischen technisierter und lebendiger Wissenschaft' und ‚Ach Du liebe Wahrheit – eine Nachbetrachtung aus dem Dialog mit Markus Stepanians' im vorliegenden Band – fortan kurz ‚*Zukünfte*' bzw. ‚*Wahrheit*'. Herzbergs erste These betrifft den Anspruch der KI-Lobby, dass es sich bei ChatGPT & Co. um gottgleiche „Superintelligenzen" handelt: „Die Realität ist: Wir erschaffen Gott. [...] Es ist kein Gott im herkömmlichen Sinne, aber es ist eine Form von künstlicher Superintelligenz. [...] Bis 2029 wird die KI schlauer sein als der klügste Mensch. Und kurz darauf wird sie eine Million Mal klüger sein als der klügste Mensch" (Gawdat & Bartlett, 2021)."[1] In der Tat: Wozu bräuchte es noch Universitäten und Professor:innen, wenn man stattdessen eine gottgleiche Superintelligenz befragen könnte, deren Urteil von keinerlei menschlichen Interessen verzerrt wäre, deren Fachkompetenz universal ist und deren Fähigkeit zur Wahrheitserkennung übermenschlich wäre? Die wahre Bedrohung für die Universitäten besteht meines Erachtens darin, dass zu viele Menschen dieser Propaganda der KI-Lobby mit ihrem haltlosen Gerede von ‚Superintelligenzen' etc. auf den Leim gehen. Dann würde Künstliche Intelligenz auf natürliche Dummheit oder zumindest Naivität treffen, und das Resultat wäre nicht nur für die Universitäten verheerend. Herzbergs Destruktion dieses Mythos durch die Klarstellung, dass Textproduktionsmaschinen wie ChatGPT keine „Wahrheitsapparate" sind und es „konstruktionsbedingt" auch nicht sein können, scheint mir vor diesem Hintergrund von größter Bedeutung zu sein. Hier besteht mein Beitrag allein darin, diesen wichtigen Punkt in meinen eigenen Worten so klar wie möglich zu reformulieren. Herzbergs zweite These betrifft die Rolle der Wahrheit in den Wissenschaften und ihren Instituten. Wenn ich ihn richtig verstehe, bezweifelt er die These, dass die Universitäten Institution sind in denen, wie Hannah Arendt sagt, „seit eh und je Wahrheit und Wahrhaftigkeit die ausschlaggebenden Kriterien sind, nach denen alles, was in ihnen vorgeht, sich zu richten hat" (Arendt 2017, S. 89). Herzberg gibt zu bedenken, dass Wissenschaft auch anwendungsorientiert zu sein hat und menschlichen Bedürfnissen dienstbar gemacht werden muss. Hier gehe es nicht um die Erkenntnis von Wahrheit, sondern um die Erkenntnis des *Nützlichen*,

[1] Mo Gawdat in einem Interview mit dem Podcaster Steve Bartlett auf seinem YouTube-Kanal: ‚The Former Google X Chief Officer Mo Gadwat', Oktober 2021. – Es ist in diesem Zusammenhang heilsam und ernüchternd, sich an eine bestechend ähnliche Prophezeiung von Marvin Minsky, einen der Gründer des ‚MIT Artificial Intelligence Laboratory' in Cambridge, Massachusetts, vor einem halben Jahrhundert zu erinnern. In einem Interview mit ‚Life Magazine' vom November 1970 sagte Minsky: „In drei bis fünf Jahren [!] wird es eine Maschine geben, die über die allgemeine Intelligenz eines durchschnittlichen Menschen verfügt. […] Ich meine eine Maschine, die in der Lage sein wird, Shakespeare zu lesen, ein Auto zu schmieren, Bürointrigen anzuzetteln, einen Witz zu erzählen und sich zu streiten. Von diesem Moment an wird die Maschine beginnen, sich mit unglaublicher Geschwindigkeit weiterzubilden. In wenigen Monaten wird sie das Niveau eines Genies erreicht haben und wenige Monate später werden ihre Fähigkeiten unvorstellbar sein" (meine Hervorhebung, meine Übersetzung).

d. h. dessen, was für bestimmte (menschliche) Zwecke *effektiv* ist. In Herzbergs Augen haben wir in der theorieorientierten Erkenntnis der Wahrheit und der anwendungsorientierten Erkenntnis des Nützlichen zwei unterschiedliche, aber gleichrangige Begriffe von Wissen oder Erkenntnis. Wir hätten es hier mit einem „anderen Erkenntnisbegriff, […] dem der Effektivität" (*Wahrheit*, S. 92) zu tun. Wäre dies richtig, dann müsste die These vom zentralen Wert der Wahrheit eingeschränkt werden: Nur die theoretische Forschung würde sich an der Wahrheit orientieren, während für die anwendungsorientierten Wissenschaften der Wert der Nützlichkeit oder Effektivität zentral wäre.

These 1: „Beim Gebrauch von Textproduktionsmaschinen ist die Wahrheit rasch eines ihrer ersten Opfer, denn diese Maschinen sind konstruktionsbedingt keine Wahrheitsapparate." (*Wahrheit*, S. 94) – Mit Blick auf die Frage unseres Symposiums nach der Rolle der Universitäten „in Zeiten von Künstlicher Intelligenz" betont Herzberg, dass der Gebrauch von KI in den Wissenschaften und an Hochschulen für sich genommen nicht das Problem sei: „Was derzeit das Hochschulsystem erschüttert oder zumindest erheblich irritiert, ist nicht KI im Allgemeinen, sondern es sind die Textproduktionsmaschinen im Besonderen" (*Zukünfte*, S. 87). Das Irritierende an textgenerativen Maschinen wie ChatGPT, Gemini, Grok, Claude etc. ist, dass sie Texte in einer natürlichen Sprache wie Englisch, Deutsch oder Chinesisch generieren können, die den Texten menschlicher Autor:innen syntaktisch und stilistisch verblüffend ähneln. Aufgrund dieser syntaktischen Fähigkeiten können wir dem Anschein nach mit ChatGPT & Co. in Dialoge eintreten, die bislang allein Menschen vorbehalten waren. Vor allem können wir diese Maschinen, so scheint es, als Informationsquelle gebrauchen. Angenommen, wir wollten wissen, wer der erste Mensch auf dem Mond war, und fragen ChatGPT: ‚Wer war der erste Mensch auf dem Mond?'. Prompt erhalten wir die Antwort: ‚Neil Armstrong war der erste Mensch auf dem Mond'. Äußerlich ist dieser sprachliche Austausch ununterscheidbar von einem echten Dialog mit einer Person, die Deutsch versteht und sich in der Geschichte der Raumfahrt auskennt. ChatGPTs Antwortsatz ‚Neil Armstrong war der erste Mensch auf dem Mond' ist syntaktisch wohlgeformt und stilistisch makellos. Von größerer Bedeutung ist jedoch, dass er sich auch semantisch, d. h. seiner Bedeutung nach als wahre Behauptung über die Welt präsentiert. In Herzbergs Worten: „[D]ie Frage wird dem Anschein nach beantwortet." Hinzu kommt, dass es sich um die Behauptung einer angeblichen ‚Superintelligenz' handelt, die sich auf eine enorme Menge von Informationen stützt. Als solche genießt sie hohe epistemische Autorität.

Wohlgemerkt: „dem Anschein nach"! Denn der Schein trügt: „Textproduktionsmaschinen entgehen der Verwicklung der Sprache mit der Welt." Eine von Herzbergs zentralen Thesen ist, dass es sich bei ChatGPT & Co. um rein syntaktische Strukturmodelle ohne „semantische Referenz auf äußere Dinge" handelt (*Zukünfte*, S. 87). Textproduktionsmaschinen produzieren wohlgeformte Satzfolgen ohne „außerrechnerische[.] Weltbezüge" (*Wahrheit*, S. 93). Daraus folgt für unseren Beispieldialog, dass der von ChatGPT generierte deutsche Satz *keine* Antwort auf unsere Frage war und nicht sein konnte. Denn unsere Frage bezog sich auf die Welt der Gegenstände (im weitesten Sinne) und Eigenschaften. Konkret ging es uns um den Träger der Eigenschaft, der erste Mensch auf dem Mond

gewesen zu sein. Um sie zu beantworten, muss man die Aufmerksamkeit auf die Welt richten und herausfinden, welchem Gegenstand die durch das sprachliche Prädikat ‚… ist der erste Mensch auf dem Mond' bezeichnete Eigenschaft zukommt. Aber dieser Weg ist einem rein syntaktischen Strukturmodell wie ChatGPT versperrt, weil es über die Fähigkeit zum „außersprachlichen Weltbezug" nicht verfügt. ChatGPT besitzt nicht die semantische Fähigkeit von kompetenten Sprechern des Deutschen, sich mit seinen Äußerungen auf Dinge in der Welt zu beziehen. Die ‚Antwort' von ChatGPT enthält den *Namen* ‚Neil Armstrong', aber sie handelt nicht von Neil Armstrong. Also wird mit ihr auch nicht behauptet, dass Neil Armstrong die Eigenschaft zukommt, der erste Mensch auf dem Mond zu sein. Daraus folgt auch, dass ChatGPTs Antwort ‚Neil Armstrong war der erste Mensch auf dem Mond' keine Wahrheit ausdrückt und keine epistemische Autorität beanspruchen kann. Diese Schlussfolgerung wird auch durch epistemische Überlegungen gestützt: Wie beantwortet man die Frage, wer der erste Mensch auf dem Mond war? Anders formuliert: Was ist die korrekte Methode der Verifikation von Sätzen der Form ‚NN war der erste Mensch auf dem Mond'? Ich hatte es schon gesagt: Um diese Frage zu beantworten, muss man seine Aufmerksamkeit auf die Welt richten und herausfinden, welchem Menschen die durch das sprachliche Prädikat ‚… ist der erste Mensch auf dem Mond' bezeichnete Eigenschaft zukommt. Dazu ist ChatGPT, wie wir sahen, prinzipiell und konstruktionsbedingt nicht in der Lage. Wie also gelangt ein reines Strukturmodell, das nur Wortbausteine und syntaktische Kompositionsregeln kennt, zu seiner Antwort? Indem es unsere ursprüngliche Frage nach dem Träger einer bestimmten Eigenschaft in der Welt durch eine andere, innersprachliche Frage ersetzt. Vereinfacht gesagt, lautet die sinnverschiedene Ersatzfrage, die ChatGPT uns heimlich unterschiebt: ‚Was ist der wahrscheinlichste Eigenname, der das deutsche Prädikat ‚… ist der erste Mensch auf dem Mond' zu einem syntaktisch wohlgeformten Satz der deutschen Sprache ergänzt?' So verstanden wird hier nicht nach dem Träger einer Eigenschaft in der außersprachlichen Welt gefragt, sondern nach etwas Sprachlichem. In ChatGPTs Ersatzfrage geht es um die statistische Wahrscheinlichkeit des Vorkommens eines *Eigennamens* an einer bestimmten Stelle in einem deutschen Satz. Mit Blick auf seinen enormen Erfahrungsschatz in Gestalt der Unmengen von Texten, mit denen ChatGPT trainiert wurde, lautet die Antwort: Der wahrscheinlichste, statistisch zu erwartende Eigenname ist ‚Neil Armstrong'. Das mag richtig sein, ist aber keine Antwort auf unsere ursprüngliche Frage. Bestenfalls erfahren wir hier etwas darüber, welcher Satz über den ersten Menschen auf dem Mond kompetente Autor:innen im Jahr 2026 gemeinhin *für wahr halten*. Aber: „Wahrsein ist etwas anderes als Fürwahrgehaltenwerden, sei es von einem, sei es von Vielen, sei es von Allen, und ist in keiner Weise darauf zurückzuführen. Es ist kein Widerspruch, dass etwas wahr ist, was von Allen für falsch gehalten wird." (Frege 1893, XV f.). In der Schule würde ChatGPT mit seiner heimlichen Ersetzung einer Sachfrage über die Welt durch ein syntaktisches Ergänzungsproblem nicht durchkommen. Dort würde es kühl heißen: ‚Thema verfehlt, Note: ungenügend'.

Wir dürfen uns also von der syntaktischen und stilistischen Virtuosität dieser Maschinen nicht blenden lassen: „Klug erscheinende Antworten sind darin genauso angelegt wie das Phantasma und der Unsinn. Wer hier nach Intelligenz oder Wahrheit sucht oder beides einzubauen hofft, findet dazu keine Ansatzpunkte. Wahrheit gehört nicht zu den

konstitutiven Kriterien, nach denen ein Sprachmodell bei der Textproduktion operiert." (*Wahrheit*, S. 93). Der gegenteilige Eindruck von Intelligenz und Verständnis beruht auf einer kognitiv-hermeneutischen Täuschung, die Joseph Weizenbaum schon 1966 beschrieben hat. Weizenbaum hatte Mitte der 1960er ein primitives Computerprogramm namens ELIZA geschrieben, das auf sehr primitive und mechanisch-schablonenhafte Weise die Interviewtechnik des Gesprächstherapeuten Carl Rogers parodieren sollte. Aber zu Weizenbaums Überraschung – und zunehmender Bestürzung – genügte schon diese schablonenhafte Parodie, um bei einigen Testpersonen eine hartnäckige Illusion zu erzeugen. Wie Weizenbaum berichtet, waren einige Probanden nur „sehr schwer davon zu überzeugen, dass ELIZA … *kein* Mensch ist" (Weizenbaum 1966, S. 42; meine Übers.). Das ist der ‚ELIZA-Effekt'. Er beruht auf der hermeneutischen Notwendigkeit, die Autor:innen sprachlicher Äußerungen, die wir verstehen, als rationale Wesen mit einem komplexen geistigen Leben aufzufassen. Die Zuschreibung einer Vielzahl (mehr oder weniger) rationaler psychischer Einstellungen in Form von Meinungen, Wünschen und kommunikativen Absichten ist eine Bedingung sprachlichen Verstehens.[2] Seitens verständiger Hörer und Leser erfolgt diese Zuschreibung unwillkürlich und instinktiv. Sie wird durch eine psychische Disposition ausgelöst, die kompetenten Sprecher:innen zur zweiten Natur geworden ist. Wir können diese Neigung selbst dann nicht unterdrücken, wenn wir wissen, dass wir es nicht mit den Äußerungen eines sprachlich kompetenten und rationalen Wesens zu tun haben, sondern z. B. mit denen eines Papageis, der Wörter, Phrasen oder Sätze nur verständnislos nachplappert.[3] Die afrikanischen Graupapageien eines Zoos in England, die vor einigen Jahren als ‚curse gang' Schlagzeilen machten und den Zoodirektor regelmäßig und scheinbar unflätig mit /jɔːr ə fæt twæt/ begrüßen, sprechen kein Englisch und beleidigen niemanden. Was sich zunächst wie eine sprachliche Äußerung mit den korrelativen psychischen Einstellungen anhört, ist in Wahrheit keine. Der weise Zoodirektor weiß, dass die putzigen Kerlchen nicht meinen, was sie zu sagen scheinen, und wird sich achselzuckend wieder seiner Arbeit zuwenden. Dennoch ist diese Reaktion die nachträgliche Korrektur einer Zuschreibung, die sich ihm unwillkürlich aufdrängt. Ähnliches gilt von den „stochastischen Papageien", die uns die KI-Lobby als gottgleiche Superintelligenzen anpreist: „Entgegen dem ersten Eindruck ist [die textgenerative KI] ein System, das Sequenzen von Sprachformen, die es in seinen umfangreichen Trainingsdaten beobachtet hat, zufällig miteinander verknüpft – basierend auf probabilistischen Informationen darüber, wie diese

[2] Die Notwendigkeit, die Autor:innen sprachlicher Äußerungen als rationale Wesen aufzufassen, bezeichnet Gadamer als ‚Axiom aller Hermeneutik' (Gadamer 1990, S. 376). Donald Davidson hat es in der modernen Sprachphilosophie als ‚Prinzip der wohlwollenden Interpretation' bekannt gemacht. In ähnlicher Weise hat der Philosoph Richard Grandy betont, dass es „pragmatische Bedingung" allen Verstehens sei, dass wir dem Autor eine Vielzahl von kognitiven und konativen Einstellungen zuschreiben, die unserem eigenen Netz von Überzeugungen und Wünschen „möglichst ähnlich" sind. Grandy bezeichnet diese interpretative Maxime als ‚Humanitätsprinzip' (Grandy 1973, S. 443). – Für einen hilfreichen Überblick mit Diskussion und Verweisen auf die einschlägige philosophische Literatur, s. Künne 1990.

[3] Wir befinden uns dann in einer ähnlichen Lage wie eine Person, die mit einer Müller-Lyerschen Streckentäuschung konfrontiert ist. Obwohl ihr diese berühmte optische Täuschung wohlvertraut sein mag und sie also weiß, dass die zwei Pfeile exakt gleichlang sind, kann sie nicht umhin, sie als verschieden lang wahrzunehmen. Sie weiß, dass sie dem Zeugnis ihrer Sinne in diesem Fall nicht trauen sollte, und kann sich dennoch des visuellen Eindrucks der Längenverschiedenheit nicht erwehren.

kombiniert werden können, jedoch ohne Bezug zur Bedeutung: ein stochastischer Papagei.“ (Bender & al., 2021, S. 616f. – Vgl. die zugängliche Darstellung in Bender und Hanna 2025). Die angemessene Reaktion auf die marktschreierischen Slogans der KI-Lobby besteht also darin, sich achselzuckend seiner wissenschaftlichen Arbeit zuzuwenden und augenrollend etwas von ‚stochastischen Papageien‘ zu murmeln. Die stochastischen Papageien plappern, die Karawane der Wissenschaft zieht weiter.

These 2: Angewandte Wissenschaft „basiert auf einem anderen Erkenntnisbegriff, […] dem der Effektivität.“ (*Wahrheit*, S. 92) – Hinsichtlich der Beziehung von KI und Wahrheit weiß ich mich mit Herzberg also einig. Aber er hat ernste Zweifel bezüglich der These, dass Wahrheit nach wie vor den zentralen Leitwert der Universitäten bildet.[4] Herzberg führt zunächst an, dass ihm „die Wahrheit im akademischen Alltag eher in ihrem Gegenteil, der Unwahrheit“, d. h. in Gestalt von „Täuschung, Betrug, Plagiat, Fälschung, Irreführung, Desinformation und die Lüge“ (*Wahrhei*t, S. 92) begegnen würde. Aber das scheint mir meine These eher zu bestätigen. Ist Täuschung, Betrug, Plagiat etc. im akademischen Alltag nicht eben deshalb ein Problem, weil Wahrheit der akademische Grundwert ist, dem Universitäten sich qua Universitäten verpflichtet sehen? Jedenfalls scheint mir die Allgegenwart von Unwahrheit im Universitätsbetrieb meine These nicht zu falsifizieren. Der Wert der Wahrheit spielt für Universitäten eine ähnliche Rolle wie der Wert der Gesundheit für Krankenhäuser. Dass Ärzt:innen sich in ihrem medizinischen Alltag vor allem mit Krankheit konfrontiert sehen, scheint mir nicht zu widerlegen, sondern zu bestätigen, dass Gesundheit der Leitwert ist. Analog für Unwahrheit und Wahrheit an der Universität. Ein zweites, überzeugenderes Argument von Herzberg stützt sich auf die Unterscheidung von theoretischer und angewandter Wissenschaft. Der Leitwert einer anwendungsorientierten Wissenschaft wie z. B. der Informatik sei nicht Wahrheit, sondern Nützlichkeit oder Effektivität: „Praktisches Wissen entsteht aus der Beobachtung, dass ein Effekt E mit einer Handlung H erreicht werden kann. Ist das mit einer gewissen Zuverlässigkeit der Fall, liegt ein Wissen um den Zusammenhang E per H vor; man nennt dieses Wissen auch eine Regel. *Eine solche Regel muss nicht wahr sein, sie muss sich lediglich als effektiv erweisen*“ (*Wahrheit*, S. 92; meine Hervorh.). Verständigen wir uns zunächst über den Wissensbegriff! Wissen ist nicht dasselbe wie Gewissheit, denn Wissen impliziert Wahrheit. Wir mögen uns beispielsweise völlig gewiss sein, dass Buzz Aldrin der erste Mensch auf dem Mond war. Gleichwohl wissen wir es nicht, da es falsch ist. Wahrheit ist eine notwendige Bedingung für Wissen. Man kann nicht wissen, dass Aldrin der erste Mensch auf dem Mond war, man kann es nur zu wissen glauben. Allgemein gilt: Wenn ich weiß, dass etwas der Fall ist, dann ist es wahr. Das gilt auch für Wissen über Kausalzusammenhänge wie in Herzbergs Beispiel. Wenn ich weiß, dass Handlungen vom Typ H mit einer gewissen Zuverlässigkeit den Effekt E hervorbringen, dann muss es wahr sein, dass sich auf H-Handlungen in der Regel ein E-Effekt einstellt. Entgegen dem, was Herzberg mit dem hervorgehobenen Satz im obigen Zitat sagt, muss die Regel, die diesen Wahrscheinlichkeitsnexus formuliert, wahr sein. Es ist ungereimt, zu sagen: ‚Ich verfüge über das praktische Wissen,

[4] Ich sollte vielleicht klarstellen, dass ‚Universität‘ hier für alle Institutionen steht, in denen Wissenschaft betrieben wird, also Technische Hochschulen und Fachhochschulen eingeschlossen.

dass der Effekt E regelmäßig auf Handlung H folgt – aber ich weiß nicht, ob das wahr ist.‘ Alles Wissen, gleichgültig, ob es praktisches oder theoretisches Wissen ist, impliziert begrifflich Wahrheit. Auch Formulierungen von praktischem Wissen in Behauptungen über Effektivitäts- oder Nützlichkeitsbeziehungen erheben einen Wahrheitsanspruch; und die Wahrheit, die sie beanspruchen, ist in beiden Fällen dieselbe. Mehr noch, Wahrheit hat hier den Primat. Wahrheit und Effektivität sind keine gleichrangigen Werte, weil Effektivität gegenüber Wahrheit parasitär ist. Auch die Effektivität von Herzbergs Regel ‚Wenn H, dann wahrscheinlich E‘ ist in ihrer Wahrheit fundiert. Die Regel ist effektiv, *weil* sie wahr ist. Der statistische Kausalnexus besteht, weil die Welt so ist, wie sie ist. Auch Wahrscheinlichkeitsaussagen sind Aussagen über das So-Sein der Welt. Alle Anwendungen sind Anwendungen von Wahrheiten, oder zumindest dessen, was man dafür hält. Auch die Landvermesser des alten Ägypten, die sich nach jeder Flut des Nils an eine Neuvermessung der Ackergrenzen machten, konnten diese praktische Aufgabe nur erfüllen, weil sie über weitreichendes theoretisches Wissen über geometrische Wahrheiten verfügten. Ob sie sich darüber im Klaren waren oder nicht: Ihr praktisches Wissen und ihre Fähigkeiten bei der korrekten Landvermessung war in solchen – theoretischen – Wahrheiten fundiert. Wahrheit ist in diesem Sinne grundlegend für menschliches Denken und Handeln. Die Metaphorik von ‚Fundament‘ und ‚Grund‘ sollten wir nicht nur in den Wissenschaften ernstnehmen. In einem schönen Bild von Hannah Arendt ist Wahrheit „der Grund, auf dem wir stehen, und der Himmel, der sich über uns erstreckt“ (Arendt 2017, S. 95).

Literatur

Arendt, Hannah: ‚Wahrheit und Politik‘ in: *Wahrheit und Lüge in der Politik – Zwei Essays*, Piper 2017, S. 44-95.

Bender, Emily M. & al.: ‚On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?’ Proceedings of the 2021 ACM on Fairness, Accountability, and Transparency, 610-623, S. 616f.

Bender, Emily M. und Hanna, Alex: *The AI Con – How to Fight Big Tech’s Hype and Create the Future We Want*, Harper Collins 2025.

Frege, Gottlob: *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Band I. Jena 1893.

Gadamer, Hans-Georg: *Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik*, Tübingen 1990.

Gawdat, Mo und Bartlett, Steven: ‚The Former Google X Chief Officer Mo Gawdat‘, Podcast vom Oktober 2021.

Weizenbaum, Joseph: ‚ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine‘, in: *Communications of the ACM*, Band 9, Nummer 1, S. 36–45.

Institutionelle Verantwortung in der Wissenschaft und in der KI-Nutzung – Fragen ans Bildungsverständnis

Ines Langemeyer

Wenn wir in diesen Zeiten über die Universität, die wissenschaftliche Bildung und über ihre Zukunft sprechen, stehen wir vor den Herausforderungen der Technologie der „Künstlichen Intelligenz". Diese sind aus bildungswissenschaftlicher Sicht keine technischen, so als wären bestimmte Funktionen nur noch zu verbessern (z.B., weil Chatbots zum „Halluzinieren" tendieren statt korrekterweise Nicht-Wissen zu artikulieren). Sie betreffen ungeachtet der Funktionsmängel das institutionelle Verständnis von Wissenschaft und Universität. Die neuen Handlungsmöglichkeiten, die wir mit KI-Anwendungen sicherlich haben bzw. entwickeln können, und ihre Folgen müssen auf einer Metaebene betrachtet werden. Wir reflektieren mithin den Umstand, wie kognitive und emotionale menschliche Fähigkeiten von Maschinen imitiert und immer perfekter simuliert werden, so dass sie sich sogar als Gesprächspartner für persönliche Probleme anbieten (vgl. Langemeyer et al., 2025). Hier, in der Entkopplung der technischen Informationsverarbeitung von menschlich geteilten und kooperativ entstehenden Lebensbezügen, steckt ein grundlegendes Problem. Technische Prozesse erzeugen nun Bedeutungen und dieses Produkt verselbstständigt sich von menschlicher Subjektivität. Das Eigenleben der Technologie (obwohl es kein Leben darstellt) strahlt aus auf Institutionen wie die Universität, auf die Menschen vertrauen, und zwar in besonderer Hinsicht, da durch sie wissenschaftliches Wissen entstehen und akademische Bildung garantiert werden soll.

Ganz generell verändern sich mit KI zum einen das Herstellen und Verfügbarmachen von Informationen, zum anderen delegieren Menschen an Large-Language-Models mehr und mehr die Erzeugung, Prüfung, Legitimierung und die interpretierende Perspektivierung und Nutzung von Daten, Verfahren und Urteilen. So sind die Bildung und die Sozialität der Menschen in einer Weise von KI betroffen, die unsere gesellschaftliche und auf uns selbst rückwirkende Subjektivität verändert. Auch die Art, wie sich gesellschaftliche Akteure sehen, interpretieren und wie sie sich gegenseitig wahrnehmen und beobachten, unterliegt einer tiefgreifenden Veränderung.

Gerichtsurteile spiegeln dieses Problem zum Beispiel in prüfungsrechtlichen Streitfällen wider, indem sie bei der einmaligen KI-Nutzung schon einen Täuschungsversuch des Prüflings erkennen, während ein Plagiat erst dann zur widerrechtlichen Täuschung wird, wenn eine Prüfungs- oder Studienleistung insgesamt vom Plagiieren geprägt ist (Zehnthöfer, 2026). Gesellschaftlich benötigen wir, so unser herkömmliches Selbstverständnis, eine gewisse Einsicht in die Eigenständigkeit von menschlichen Leistungen und Kompetenzen, insofern wir diese gewissenhaft (und nicht leichtfertig) einer Person zurechnen und uns auf dieses Urteil auch längerfristig verlassen wollen. Dies macht den institutionellen Charakter von Bildung aus.

Sowohl die Lernerfahrung als auch die gemeinsame, miteinander geteilte Erfahrung werden aber durch die Art des Synthetisierens und Verfügbarmachens von Informationen

mittels KI verkürzt und entkontextualisiert. Wir verlieren ein Verständnis von Leistung und Kompetenz und wir verlieren stückweise unsere Kompetenz, die Leistungen und Kompetenzen wechselseitig richtig einzuschätzen, ohne weitere Hilfsmittel wie einen Chatbot oder eine Suchmaschine zu benutzen. Wir werden dazu angehalten, blindes Vertrauen zu praktizieren. Wir haben zugleich noch wenig Erfahrung damit, wie vertrauenswürdig KI überhaupt ist, weil sie von ganz unterschiedlichen Herstellern bereitgestellt werden kann. Eigentlich müssten wir plötzlich Misstrauen haben in all das, was wir hier konsumieren können. Aber das Nicht-Wissen über die Art der Leistung und der Verfahrensweise der Algorithmen schützt uns davor.

Warum und inwiefern muss man einen solchen Verlust von Wissen und das blind praktizierte Vertrauen beklagen und genauer reflektieren? In einem direkten zwischenmenschlichen Kontakt setzen wir die Subjektivität bei anderen auch oft stillschweigend voraus. Im mündlichen Austausch mit körperlicher Anwesenheit erlernen wir jedoch einzuschätzen, vor welchem Hintergrund jemand zu uns spricht und argumentiert. Wir fragen uns, wenn jemand z.B. eine Tatsachenbehauptung aufstellt, was die Quelle seines Wissens ist, inwiefern er eine bestimmte Verantwortung trägt oder eine besondere Kompetenz besitzt – oder ob wir sagen müssen, dass sich jemand einmischt, um sich wichtig zu machen oder um persönliche Vorteile zu erlangen. Solche metakognitiven Prüfungen werden oftmals eher beiläufig und unterstützt durch eine emotionale Ebene geleistet. Ohne es ganz genau besser zu wissen als unser Gegenüber, entdecken wir körperliche Anzeichen, die uns z.B. an der Vertrauenswürdigkeit zweifeln lassen. Es ist das, was die implizite Seite unserer Wahrnehmung ausmacht. Im Spüren der Leiblichkeit von uns selbst und anderen entsteht über den vielfältigen Kontakt im Leben mit Menschen und durch die Unterstützung unserer Kritikfähigkeit, wie sie zum heutigen Verständnis von Allgemeinbildung gehört, eine individuelle Erfahrungsgrundlage, was es heißt, kritisch zu sein. Vor dem Hintergrund dieser Erfahrung vervielfältigen sich Formen, wie wir auf das (sprachliche) Handeln von anderen sinnvoll mit eigenen (sprachlichen) Handlungen reagieren oder wie wir ihre Kompetenz und ihre Anliegen und Bedürfnisse einschätzen können (Langemeyer et al., 2025; Langemeyer et al., 2026). Dies erstreckt sich auch weiter auf die unmittelbaren und nicht mehr unmittelbar gemachten Erfahrungen in und mit Institutionen – und die wissenschaftliche Bildung ist eine davon. Auch wissenschaftliche Bildung, wissenschaftliches Wissen und die Universität werden mit einem bestimmten Vertrauen oder Misstrauen gesehen, was auf die Handlungsmöglichkeiten dieser Institutionen rückwirkt.

Es bestand somit ein Problem moderner Gesellschaften bereits vor der Verbreitung von KI-Anwendungen, inwiefern ihre Institutionen vertrauensvoll sein können und wie sie Vertrauen und Verlässlichkeit herstellen. Die Universität bzw. die Wissenschaft befriedigt ein Bedürfnis nach Vertrauenswürdigkeit in wissenschaftliche Expertisen, indem sie innerhalb ihrer disziplinspezifischen Fachbereiche eine innerdisziplinäre Kontrolle initiiert, koordiniert und vorschreibt. Sie organisiert zudem ein Prüfungswesen.

Doch was diese Kontrolle ausmacht, verändert die KI. Die institutionelle und erfahrungsgeschichtliche Einbettung von Sprache und Wissen in unseren Lebenszusammenhang wird in den Such-, Analyse- und Synthetisierungsfunktionen von KI übersprungen. Damit

werden generationenübergreifende Lernerfahrungen begradigt und entkernt. Informationen digital zu teilen, wird ein vorwiegend technisches Verfügbarmachen und es verschwindet eine gelebte Form von Erfahrungsaustausch, der nicht nur in der Aufnahme oder Weitergabe von Informationen besteht (Langemeyer et al., 2022). Im sprachlichen Austausch werden immer weniger spontane und institutionell verantwortete Korrekturen mit den instant konsumierbaren Informationshäppchen verbunden. Ein gesellschaftlicher Konsens über dieses „Wissen" wird technisch vorausgesetzt, nicht durch den Austausch errungen.

In den Chatbots tritt das technische Verfügbarmachen zugleich in einer perfektionierten Sprachgestalt in Erscheinung, so dass Vertrauen in diese Technologie aus einem recht trivialen Grund besteht. Dennoch fehlt einer oberflächlichen Erscheinung von Expertise in vielfacher Hinsicht der Hintergrund der menschlichen Erfahrung und der menschlichen Art, Vertrauen aufgrund von Kritikfähigkeit und Irrtumskorrekturen aufzubauen und Verantwortung zu übernehmen (Langemeyer, 2024; Langemeyer & Ebner von Eschenbach, 2026). Wie Kurt Möller (2022, S. 20) mit Bezug auf John Dewey schreibt, betrifft diese Verantwortungsfrage auch das Demokratieverständnis in ideeller Form:

„Demokratie [ist] nicht nur eine Regierungsform […], sondern eine Form der gemeinsamen und miteinander geteilten Erfahrung sowie ein vor allem durch die Ideen von Gleichheit und Freiheit normativ aufgeladenes Prinzip sozialer Assoziation und Kooperation."

Wissenschaft nimmt diese Prinzipien von Gleichheit und Freiheit in sich auf, wenn es ihr darum geht, dass nur solchen Urteilen und nur solchem Wissen Vertrauen zu schenken ist, das von den kompetenten Erkenntnissubjekten unter Anstrengungen ihrer Vernunft ohne Rücksicht auf Machtstellung und Ansehen geprüft wurde. Wissenschaftliche Erkenntnissuche richtet sich in ihrem Anspruch gegen ein Geheimwissen, gegen Privilegien der Vorteilsnahme und gegen eine standesmäßige oder autoritäre Entscheidungsmacht über Wahrheit; wie der jüngst verstorbene Philosoph Jürgen Habermas sagte, soll ein besseres Argument zwanglos zu einem „Zwang" werden, d.h. aufgrund von vernünftiger Einsicht zur allgemeinen Geltung gelangen. Wissenschaft teilt insofern mit dem ideellen Selbstverständnis heutiger Demokratie die Grundzüge von Gleichheit und Freiheit als allgemeine Prinzipien unseres Handelns. Insofern richten sich die gesellschaftlichen Erwartungen an die Realisierung dieser Prinzipien an die Institutionen der Wissenschaft.

Ein KI-Tool ist aber keine Person und noch längst keine Institution geworden. Sie ist nicht verantwortlich und auch nicht von Gleichheit und Freiheit betroffen. Denn eine KI kann weder Einsicht nehmen in Erfahrungen von Ungleichheit und Unfreiheit noch sich im Hinblick auf die Werte von Gleichheit und Freiheit selbst reflektieren. Daher trägt diese Technologie niemals Verantwortung.

Von „kognitiven Schulden" oder „Skill-Skipping" wird mittlerweile in Studien in Bezug auf die Nutzer gesprochen, weil die technologischen Möglichkeiten bestehen, statt selbst nach Zusammenhängen oder Lösungen zu suchen, solche Aufgaben im jeweiligen Moment des Auftretens an die Maschinen zu delegieren. Datenstrukturen, die aus Nutzungsdaten im alltäglichen Leben generiert werden, spielen dabei eine bedeutende Rolle. Auf der Grundlage von Graphen erzeugt KI eine eigene semantische Welt von Beziehungen, die

Maschinen mimetisch (Loheit, 2025) und ordnungsstiftend werden lässt. Dabei spielt die Anwendung der Graphentheorie eine wichtige Rolle (vgl. Dath, 2026). Die ordnungsbildenden Graphen werden zwar ständig mit neuen Daten konfrontiert. Sie sind somit adaptiv, indem sich neue Kanten und Kurven zwischen den Knotenpunkten sowie neue Knotenpunkte bilden können. Mit Dietmar Dath ist jedoch zu fragen, inwiefern es noch eine Rücksicht aufs Besondere gibt, wie es sich auch in unseren Bedürfnissen und Bedürfniskonflikten als Komplexität menschlichen Lebens zeigt. Anders gesagt, wenn es zur Unterstützung unserer Aktivitäten durch eine KI eine technische Adaption auf einen gesellschaftlichen Status Quo oder auf das eigene subjektive Präferenzverhalten gibt, was hilft uns dann zu erkennen, wer wir sind und warum wir die sind, die wir sind? Können menschliche Beziehungen hier tatsächlich technisch ersetzt werden?

Während menschliche Intelligenz und das sozial geteilte Erleben unserer Lebenssituationen es uns ermöglichen, einer permanenten Ausdifferenzierung, Zerlegung und Auffächerung eines semantischen Gefüges von Bedeutungen entgegenzuwirken, gleichzeitig eigene Perspektiven und Autonomie darin zu gewinnen, bauen sich die neuen Maschinen in solche semantischen Geflechte ein, wo sie uns im Moment der Unterstützung, wenn wir sie als Helfer nehmen, zugleich darin behindern, uns und andere als Verantwortliche zu erleben und den Zusammenhang der Verantwortung für unser Leben reflexiv zu erfahren. Selbstbewusstsein entsteht jedoch damit, dass wir uns für etwas verantwortlich fühlen. Mit der Übernahme der Verantwortung in einem Kontext fangen wir an, uns bewusst zu reflektieren. Hier verbinden sich bewusst gemachte Bedeutungen mit gelebten Beziehungen in einem Lebenskontext. Deshalb sind Universitäten nicht nur Orte der Forschung, sondern auch Orte der Bildung. Sie verwirklichen ihren institutionellen Charakter im Einzelnen, im Bildungssubjekt. Und dies geschieht umso mehr, je mehr ihre Mitglieder etwas mit der Verantwortung für Wissenschaft und einer wissenschaftlichen Bildung anzufangen wissen.

Wir geben diese Orte der institutionellen Reproduktion von Wissenschaft allerdings preis, wenn der Zusammenhang der Verantwortungsübernahme und der -reflexion leidet, wenn die Kritik und die Korrektur an der Wissensvermittlung durch die technische Apparatur des Verfügbarmachens den Subjekten entzogen wird. Bedeutungen lösen sich in Strukturen auf, die „Künstliche Intelligenz“ u.a. mittels Graphen technisch erzeugt. Wir können uns an Günther Anders‘ Kritik erinnern, die vorführt, dass die Welt medientechnisch zum Phantom und zur Matrize wird, weil Erfahrung letztlich weglos entsteht: Die Fernsehnachrichten liefern die Bilder von unterschiedlichsten Orten und geschichtlichen Ereignissen direkt ins Wohnzimmer (Anders, 1954/2018). Diese Einsicht ist noch immer triftig. Aber es ist nicht nur die persönliche Erfahrung vom Medialen betroffen, sondern auch eine überindividuelle Erfahrung, die institutionell entsteht. Eine solche Technologiekritik an den institutionellen Veränderungen von Wissenschaft und Bildung ist nötig und keine blinde Technologiefeindlichkeit.

Literatur

Anders, G. (2018) [1956]. Die Antiquiertheit des Menschen 1. Über die Seele im Zeitalter der zweiten industriellen Revolution. 4. Aufl. München: C. H. Beck.

Dath, D. (2026). Sind wir alle falsch verbunden? Frankfurter Allgemeine Zeitung, 23.04.2026

Langemeyer, I. (2024). Vertrauensvolle KI – eine Diskussion aus psychologischer und pädagogischer Sicht. In: Schork, S. (Hg.) Vertrauen in Künstliche Intelligenz. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-43816-6_6

Langemeyer, I., Strohschein, J., & Schlindwein, N. (2022). Gelingen Diskussionen im digitalen Studium? – Über die Bedeutung von Gesprächen für das wissenschaftliche Lernen. Zeitschrift für Hochschulentwicklung, 17(3), 53.

Langemeyer, I., Stockmann, F.W., Schlindwein, N. (2025). Eine transdisziplinäre Exploration zum Umgang mit KI-Chatbots im Studienalltag im Kontext von Einsamkeit–Emotionale Selbstregulation als Dimension von Bildung. Qualität in der Wissenschaft (QiW). WeblerUniversitätsverlag

Langemeyer, I.; Ebner von Eschenbach, M. (2026). „Ohne großen Respekt vor der Schulwahrheit" – Einleitung zum Band: Gaston Bachelard: Pädagogische Lektüren. In: dies. Und Ortfried Schäffter (Hrsg.). Gaston Bachelard: Pädagogische Lektüren. Springer.

Loheit, J.; Schütt, M. (2025). Mimetische Maschinen. Sozialkommunikation und Künstliche Intelligenz. Das Argument 343, S. 449-465.

Möller, K. (2022). Demokratische Bildung – Grundlegende Begriffe und Verständnisse. In: Felix Steinbrenner, Florian Neuscheler, Kurt Möller (Hrsg.). Demokratie gestalten. Herausforderungen und Ansätze für Bildungs- und Sozialarbeit, S. 16-28. Kohlhammere

Zehnthöfer, J. (2026). Es reicht der Anschein von KI. Frankfurtr Allgemeine Zeitung, 23.04.2026.

Der vergessene Faktor Zeit: Der universitäre Paradigmenwechsel von der Bildung zur Ausbildung und die Rolle von KI

Dorothea Winter

Einleitung

Die Positionen, mit denen Hochschulen, Akkreditierungsagenturen und Wissenschaftsräte seit 2023 auf generative Künstliche Intelligenz (KI) reagieren, lesen sich auffallend gleichförmig. Sie raten zur verantwortungsvollen Integration, fordern Leitlinien, betonen Chancen, mahnen Risiken an und empfehlen die Anpassung der Prüfungsformate, etwa das Abschaffen schriftlicher Haus- und Abschlussarbeiten (Hochschule Bayern 2025; Friedrich, Tobor und Wan 2024; Wissenschaftsrat 2022; Miao und Holmes 2023). Was in dieser Diskussion auffallend selten angesprochen wird, ist der bildungstheoretische Paradigmenwechsel, innerhalb dessen KI überhaupt ihre Wirkung entfachen kann. Der vorliegende Essay vertritt vor diesem Hintergrund die These, dass diese genannten Wirkungsanalysen zu kurz greifen. Ja, KI verändert, wie an Hochschulen gelehrt, geforscht und publiziert wird; allerdings findet bereits seit einigen Jahren eine Neuausrichtung der Hochschulbildung statt, bei der sich die Universität von einer Institution im Sinne einer humanistischen Bildung zu einer Institution der marktgerechten Ausbildung wandelt. KI wirkt bei diesem Prozess als Katalysator, weil sie sowohl quantitativ als auch qualitativ diese Transformation maßgeblich beeinflusst – inwiefern soll in diesem Essay analysiert werden.

1 Bildung versus Ausbildung

Spätestens seit Wilhelm Humboldts Gedanken zum bildungspolitischen Ideal universitärer Ausbildung hat sich an deutschen und später auch internationalen, daran ausgerichteten Universitäten ein Bildungsbegriff etabliert, dessen Kern Humboldt selbst um 1809/10 für die Berliner Universitätsgründung formuliert hat: Das Wesen der höheren wissenschaftlichen Anstalten bestehe darin, „innerlich die objective Wissenschaft mit der subjectiven Bildung […] zu verknüpfen“ (Humboldt 1809/1903, S. 250). Wissenschaft kann in dieser Konzeption ausdrücklich nicht auf einen außerhalb ihrer selbst liegenden Zweck reduziert werden, und Bildung wird primär als Selbstbildung der studierenden Person verstanden, die ihren Wert in sich trägt.

Diesem humanistischen Wissenschaftsideal guter Universitätsbildung setzt Humboldt als kategoriale Unterscheidung die Ausbildung als Realqualifikation entgegen, die nicht zweckfrei dem Individuum, sondern primär der Gesellschaft als definierbare Tätigkeit für den Arbeitsmarkt dient. Im Rückgriff auf Humboldt lässt sich an dieser Stelle somit den Universitäten als Orten individueller Selbstbildung im humanistischen Sinne die universitäre Ausbildung als Vorbereitung für den Arbeitsmarkt entgegenstellen.

Diese Unterscheidung hat sich seit geraumer Zeit, besonders im Zuge der Bologna-Reformen, der zunehmenden Drittmittelorientierung und der Etablierung internationaler

Rankings insofern schrittweise verschoben, als das Ideal einer freien, zweckfreien Wissenschaft der Zielsetzung einer „unternehmerischen Universität“, deren Wert sich an Output bemisst (Liessmann 2006; Münch 2009), weichen muss. So hat sich beispielsweise die Technische Universität München inzwischen offensiv als Gründungsuniversität positioniert: Im Geschäftsjahr 2024 verzeichnete sie erstmals mehr als hundert technologieorientierte Ausgründungen (TUM 2025). Universitäre Exzellenz wird zunehmend in Output-Einheiten gemessen – etwa in Patenten, Spin-offs, Beschäftigungsfähigkeit, Drittmittelvolumina –, nicht an der Tiefe wissenschaftlicher Reflexion. Julian Nida-Rümelin (2014) bezeichnet diese Entwicklung als „Akademisierungswahn“ und zeigt, wie sie nicht nur die akademische Bildung trivialisiert, sondern – und das scheint zunächst überraschend – zugleich die berufliche Ausbildung systematisch entwertet. Konrad Paul Liessmann (2006) hingegen nennt das Ergebnis dieses Prozesses „Unbildung“: die Universität als Bildungsbetrieb, in dem Wissen als Ware gehandhabt und gegen seinen eigenen Wert deformiert wird.

Die Verschiebung von Bildung zu Ausbildung, von Selbstzweck zu Marktzweck, ist demnach nicht neu, erhält jedoch in dem Maße, in dem KI zunehmende Bereiche unserer Gesellschaft und damit auch der universitären Bildung durchdringt, neue Relevanz.

2 KI und der vergessene Faktor Zeit

Mit dieser Verschiebung von Bildung zu Ausbildung einher geht eine Verknappung jener Zeit, die Studierenden für ihr Studium zugemessen wird. Wo es früher die Möglichkeit einer individuellen zeitlichen Studienplanung gab, geben heute bindende Studienverlaufspläne, fest terminierte Modulprüfungen und maximale Studiendauer Länge und Ablauf des Studiums vor. Im Rahmen dieser zeitlichen Festsetzungen wirkt KI ambivalent.

Einerseits verkürzt KI Prozesse, die zuvor Tage oder Wochen in Anspruch genommen haben, auf den Zeithorizont von Minuten: Recherche, Strukturierung, Formulierung, Korrektur, Übersetzung – all diese Tätigkeiten lassen sich in erheblichem Umfang an Sprachmodelle delegieren (Bastani et al. 2025). Diese Beschleunigung verändert das Pensum dessen, was an einem Hochschultag überhaupt leistbar ist. Die soziologische Beschleunigungsdiagnose, wie Hartmut Rosa (2005) sie für moderne Gesellschaften insgesamt entwickelt hat, gewinnt im Hochschulkontext damit eine neue Dimension: Wer etwa mit KI eine Hausarbeit in zwei Stunden statt zwei Wochen fertigstellt, nutzt die gewonnene Zeit nicht für vertieftes Lesen oder akademisches Diskutieren, sondern vermutlich für das nächste Modul, den nächsten Nebenjob, das nächste Praktikum. Die Verdichtung des Lernens legt es also nahe, in der gleichen Zeit „mehr“ zu absolvieren, statt im Möglichen länger zu verweilen. Was dabei verloren geht, ist die ganz praktische Möglichkeit, ein Argument zu verwerfen, eine zweite Quelle zu konsultieren oder einer Frage länger nachzuhängen, als es die Modulplanung vorsieht.

Andererseits – und das ist der schwerwiegendere Aspekt – verschiebt KI das Urteil darüber, welche Fähigkeiten überhaupt noch in der Hochschulbildung vermittelt werden sollen. Das führt zu Fragen wie: Wozu die Studierenden noch Essays oder Hausarbeiten schreiben lassen, wenn sie das mithilfe generativer KI problemlos generieren lassen können? Ist es nicht

bereits heute absehbar, dass Absolvent:innen ihr gesamtes weiteres Berufsleben in einer Umgebung verbringen werden, in der KI-Assistenz Standard sein wird? Empirische Erhebungen aus den deutschen Hochschulen belegen, dass generative KI längst integraler Bestandteil studentischer Arbeitspraxis geworden ist (Wannemacher et al. 2025). Ist die Forderung nach eigenständigem Schreiben daher nicht längst didaktischer Anachronismus?

Mitnichten. Was sich hier in liberalem Gewand präsentiert, erweist sich bei näherem Hinsehen als fundamentaler Wertewechsel. Lesen, Schreiben oder Vortragen sollen durch KI ökonomisiert und damit prinzipieller Marktbarkeit unterworfen werden. Was vormals als Ausdruck von Individualität – etwa durch Stil, Duktus und kulturelle Prägung – galt, wird nunmehr auf Funktionalität reduziert. Dabei sind Fertigkeiten wie das eigenständige Verfassen eines guten Essays, das Verfertigen einer kohärenten Abschlussarbeit oder das Halten eines logisch stringenten Vortrags im Seminarraum konstitutive Entwicklungsschritte humanistischer Persönlichkeitsbildung. Diese Fähigkeiten sind in ihrer Bedeutung selbst dann nicht zu unterschätzen, wenn die Absolventin sie in ihrem späteren beruflichen Werdegang kaum noch in dieser Form anwenden können muss, da sie dort vermutlich auf KI zurückgreifen wird oder andere Tätigkeiten ausübt.

Es ist somit zu kurz gedacht, wenn wissenschaftliche Disziplinen, die sich dieser ökonomischen Funktionalität qua Methodik und Inhalt eher entziehen, heute ins universitäre Abseits geschoben werden und weniger Ausstattung an Lehrstühlen, weniger Drittmittel, weniger Promotionsstellen und am Ende auch weniger Studierende erhalten. Philosophische Erkenntnisse, musikalische Ausdrucksfähigkeit oder sozialwissenschaftliche Analysen lassen sich nicht auf Knopfdruck hervorbringen und reifen nur dann, wenn ihnen die dafür erforderliche Zeit zugebilligt wird – und das vor allem im Studium. Die Gefahr für geistes-, sozial-, kultur- und musikwissenschaftliche Disziplinen lautet im KI-Zeitalter: Was der Markt nicht braucht oder die KI irgendwie auch kann, muss auch nicht mehr gelehrt, unterstützt oder ausgestattet werden.

Es lässt sich an dieser Stelle sagen, dass KI in der Universitätsbildung so einzusetzen wäre, dass sie Studierende bei zeitraubenden Tätigkeiten unterstützt, die für die Persönlichkeitsbildung nicht zentral sind, und damit jene zeitlichen Fenster schafft, in denen die Studierenden ihre akademische Bildung vorantreiben können. Die Frage, die sich an dieser Stelle auftut: Welche Fähigkeiten sind es, die trotz – oder gerade wegen – KI weiterhin in der Universitätsbildung vermittelt werden müssen und welche dürfen getrost wegfallen?

3 De-Skilling als Gefahr der humanistischen Universitätsbildung

Diese Frage ist zuvörderst eine bildungstheoretische, denn die Verwendung von KI in der Universitätsbildung hat eben nicht nur die obig angesprochene zeitliche, d. i. quantitative, Dimension, sondern auch eine qualitative. In dem Maße, in dem Studierende individuelle Fertigkeiten dauerhaft an KI auslagern, reduzieren sie diese Fertigkeiten auf individueller Ebene, aber auch auf gesamtgesellschaftlicher: Indem immer mehr Einzelne weniger Fertigkeiten erwerben, verringert sich deren Summe auch gesamtgesellschaftlich; es tritt eine Verarmung ein, im schlimmsten Fall verliert die Gesellschaft diese jeweiligen Fertigkeiten ganz.

Anil Doshi und Oliver Hauser (2024) haben in dem Zusammenhang etwa gezeigt, dass der Zugang zu LLM-generierten Ideen die individuelle Kreativität von Kurzgeschichten zwar messbar steigert – insbesondere bei weniger kreativen Schreibenden –, dass die so entstehenden Texte aber zugleich einander deutlich ähnlicher werden als die rein menschlich verfassten. Die Autoren sprechen daher von einem „social dilemma" der Kreativität: individuelle Verbesserung bei kollektiver Verarmung des Neuen (Doshi und Hauser 2024). Generative KI erzeugt also messbar Mittelmaß auf gehobenem Niveau, verhindert aber zugleich jene Ausreißer, die Leuchttürme, die Universitätsbildung hervorbringen soll und die wissenschaftliche Grenzen verschieben können. Man könnte pointierter sagen: Nichts weniger als wissenschaftlicher Fortschritt per se wird dadurch bedroht.

Hinzu treten Befunde zur kognitiven Verfasstheit der Schreibenden selbst. Eine Untersuchung am MIT Media Lab dokumentierte mittels EEG und Textanalyse über vier Monate hinweg eine messbare Reduktion neuronaler Aktivierung sowie sprachlicher Eigenständigkeit bei Proband:innen, die argumentative Aufgaben mit ChatGPT-Unterstützung bearbeiteten. Die Autor:innen sprechen vorsichtig von einer „cognitive debt" und einem wahrscheinlichen Rückgang von Lernfähigkeiten bei früher Übernahme der Schreibarbeit durch das Sprachmodell (Kosmyna et al. 2025). Eine kontrollierte experimentelle Studie von Georgiou (2025) bestätigt diesen Befund psychometrisch: Proband:innen mit ChatGPT-Zugang zeigten signifikant niedrigere Werte auf einer Skala kognitiver Beteiligung als eine Kontrollgruppe ohne KI. Die zugrunde liegende Mechanik wird „cognitive offloading" genannt. Wer mentale Arbeit zuverlässig externalisieren kann, investiert weniger Aufwand in das Encoding, was die langfristige Verankerung von Wissen maßgeblich schwächt (Risko und Gilbert 2016).

Die empirischen Befunde lassen den normativen Schluss zu, dass, wenn KI als zeitlicher Beschleuniger eingesetzt wird, dies zu einem schleichenden Niveauverlust führen kann. Was für die Studierenden als Erleichterung des Studiums und Effizienzgewinn wirken mag, endet im kollektiven De-Skilling. So ist etwa die Fähigkeit, einen guten Essay zu schreiben, nicht nur funktional, sondern vielmehr konstitutiv für die Selbstführung des Denkens. Wer schreibt, zwingt sich, Verstandenes zu prüfen, Ungesagtes zu artikulieren, Lücken im eigenen Argument zu erkennen. Schon Heinrich von Kleist hat in seiner kleinen Schrift *Über die allmähliche Verfertigung der Gedanken beim Reden* (1805) gezeigt, dass das Formulieren ein Verfahren der Erkenntnisproduktion ist und nicht ihrer bloßen Wiedergabe. Damit stellt sich die Frage: Was dagegen tun?

4 KI als Enhancing-Tool für Studierende

KI aus der Universitätsbildung verbannen zu wollen, ist weder realistisch noch sinnvoll. Mithilfe von KI lassen sich, wie bereits an vorheriger Stelle betont, sehr wohl auch in der Universitätsbildung Aufgaben erledigen, deren eigenständige Aneignung keinen inhärenten Eigenwert für die Bildung der Studierenden besitzt.

Welche Aufgaben genau an KI ausgelagert und damit automatisiert werden können, müssen Bildungswissenschaftler:innen im Einzelnen untersuchen und festlegen. Aus philosophischer Perspektive lässt sich grundsätzlich wenig gegen KI als „Adlatus" – im Sinne

eines wissenschaftlichen Beistands in untergeordneter Funktion – vorbringen, der Studierenden „zuarbeitet“. So können Studierende mithilfe von KI Literaturrecherchen vorstrukturieren, Quellen in Übersicht bringen, fremdsprachige Texte erschließen, Sachverhalte auf Anforderung erläutern, Grammatik und Rechtschreibung prüfen oder terminologische Konsistenz gewährleisten.

In all diesen Fällen entlastet KI von Tätigkeiten, deren Beherrschung nicht Selbstzweck der akademischen Ausbildung ist, sondern Voraussetzung für den eigentlichen Bildungsvorgang. Wer Stunden mit der Suche nach einer Bibliothekssignatur verbringt, lernt dabei wenig, was nicht effizienter zu erreichen wäre. Wer die fünfte Korrekturschleife einer englischsprachigen Passage absolviert, ebenso.

Die universitäre KI-Verwendung lässt sich didaktisch durch Hochschullehrende operationalisieren. Prozessbegleitende Schreibportfolios, in denen einzelne Stadien des Schreibens dokumentiert werden, machen die Eigenleistung sichtbar. Mündliche Prüfungen sichern, dass die schriftlich vorgelegte Argumentation auch tatsächlich vollzogen wurde. Eine explizite curriculare Reflexion auf die Differenz zwischen menschlicher Autorschaft und maschineller Generierung (nicht als technologisches, sondern als bildungstheoretisches Thema) gehört in die Eingangsphase jedes Studiengangs.

Damit verschiebt sich auch das, was von Hochschullehrenden zu fordern ist. Wer KI aus dem Seminarraum heraushalten will, lehrt an einer Wirklichkeit vorbei, die längst zur studentischen Praxis gehört. Notwendig wird also ein Kompetenzaufbau, der über die Bedienung der Werkzeuge hinausgeht und deren eigentümliche Funktionsweise einschließt: dass Sprachmodelle stochastisch operieren, plausibel klingende Fehlinformationen produzieren und in ihren Ausgaben die Verzerrungen ihrer Trainingsdaten reproduzieren (Friedrich, Tobor und Wan 2024). Erst auf dieser Grundlage lassen sich Prüfungen sinnvoll umbauen, ohne die Eigenleistung der Studierenden aufzugeben, und erst dann ist plausibel zu fordern, dass auch Lehrende selbst offenlegen, wo sie KI in eigene Lehrmaterialien, Bewertungen oder Forschungsroutinen einfließen lassen (Hochschule Bayern 2025; Wannemacher et al. 2025). Eine humanzentrierte KI-Integration, so betont es die UNESCO, hängt weniger an der technischen Ausstattung der Hochschulen als an der Befähigung ihrer Lehrenden (Miao und Holmes 2023).

5 Der Elitismusvorwurf und die KI-Inklusionsperspektive auf Universitätsbildung

Gegen die hier vertretene Verteidigung des humanistischen Bildungsideals lässt sich natürlich ein gewichtiger Einwand erheben. Humboldts Vision sei historisch das Privileg einer schmalen, vorwiegend männlichen, bürgerlich-akademischen Schicht gewesen, deren ökonomische, kulturelle und zeitliche Ausstattung jene zweckfreie Selbstbildung überhaupt erst ermöglichte. Das Festhalten an einem Ideal, das in seiner historischen Realisierung breite Bevölkerungsteile systematisch ausschloss, gerät unter den Verdacht des Elitismus und der Exklusivität. Studierende mit Behinderung, mit nicht-deutscher Erstsprache, mit Care-Verpflichtungen, in prekärer ökonomischer Lage oder mit neurodivergentem

Profil finden in der idealtypischen Humboldt-Universität mit ihrem Vorrang langer Lesezeiten, eigenhändiger Quellenarbeit und akademisch normierter Sprachform Zugangsbarrieren, die nicht ihrem intellektuellen Vermögen, sondern den Voraussetzungen des Modells geschuldet sind. Das ist erst einmal richtig. Allerdings scheint gerade hier der KI-Einsatz ein Mehr an Inklusion, Barriere-Abbau und Intersektionalität zu ermöglichen, das ohne KI nicht ohne weiteres möglich ist.

Das bedeutet, dass gerade KI unter bestimmten Bedingungen nicht das Ende des humanistischen Ideals bedeutet, sondern dessen Zugang erweitert. So zeigt etwa eine integrative Übersicht zur Hochschulbildung, dass KI-gestützte unterstützende Technologien Zugänglichkeit, Inklusion und akademische Performanz von Studierenden mit Behinderung substantiell verbessern können, etwa durch Spracheingabe und -ausgabe, automatisierte Untertitelung, adaptive Textaufbereitung und individualisierte Rückmeldung (Dumitru et al. 2025). Die UNESCO formuliert diese Perspektive normativ als „human-centred vision" generativer KI, die menschliche Handlungsmacht, Inklusion, Geschlechtergleichheit sowie sprachliche und kulturelle Diversität zu fördern habe (Miao und Holmes 2023). KI wäre demnach nicht der Antagonist humboldtscher Bildung, sondern könnte, sorgsam eingesetzt, das humanistische Ideal von der historisch privilegierten Klientel ablösen und einer pluraleren Studierendenschaft erschließen. Der Maßstab dafür bleibt allerdings derselbe wie für die zeitökonomische Verwendung: KI darf Selbstbildung nicht ersetzen, sondern muss ihre Voraussetzungen breiter zugänglich machen.

Conclusio

Vor diesem Hintergrund lassen sich die Leitlinien einer humanistischen Universitätsbildung für das Zeitalter von KI nicht aus technologischer Anpassung ableiten. Sie müssen einer Reflexion auf jene Normen und Werte entspringen, die die universitäre Praxis in ihrem Selbst ausmachen.

Als oberstes Prinzip sollte darauf rekurriert werden, dass Universitäten keine Berufsausbildungsanstalten sind, sondern Orte, an denen sich Lehrende und Lernende der Wissenschaft als einem „noch nicht ganz aufgelöste[n] Problem" (Humboldt 1809/1903, S. 251) zweckfrei und in einem angemessenen zeitlichen Rahmen zuwenden dürfen. Eine Output-Logik, in der Exzellenz an Ausgründungszahlen oder Drittmittelvolumina gemessen wird, verfehlt diesen Ort.

Des Weiteren ist Zeit als bildungskonstitutiver Faktor maßgeblich. Wo immer KI als Adlata der Studierenden zum Einsatz kommt, kann sie jene Zeiträume eröffnen, in denen Wissenschaft stattfinden kann. Allerdings kann der KI-Einsatz auch De-Skilling-Effekte nach sich ziehen und Universitäten ihrer Wissenschaftlichkeit berauben, indem er Fähigkeiten der einzelnen Studierenden, aber auch des universitären Bereichs insgesamt aushöhlt.

Schließlich kann KI in summa dieser Eigenschaften eine ambivalente Wirkung entfalten: Ihre Möglichkeiten in der Universitätsbildung können sich positiv auf jene Studierenden auswirken, die sie so zu nutzen wissen, dass sie in ihrer individuellen Bildung vorankommen, oder negativ, indem sie jene, die unreflektiert auf KI zurückgreifen, in ihrer Bildung

hemmt. Hinzu tritt ein Inklusionspotenzial, das in der bildungspolitischen Auseinandersetzung mit KI bislang unterbelichtet bleibt.

Welche Form die Universitätsbildung im 21. Jahrhundert annehmen wird, hängt davon ab, welche Entscheidungen – nicht zuletzt bezüglich der Verwendung von KI – Universitätslehrende, Universitätsleitende und Studierende heute treffen.

Literatur

Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı und Rei Mariman. 2025. „Generative AI Without Guardrails Can Harm Learning: Evidence from High School Mathematics." *Proceedings of the National Academy of Sciences* 122 (26): e2422633122. https://doi.org/10.1073/pnas.2422633122.

Doshi, Anil R. und Oliver P. Hauser. 2024. „Generative AI Enhances Individual Creativity but Reduces the Collective Diversity of Novel Content." *Science Advances* 10 (28): eadn5290. https://doi.org/10.1126/sciadv.adn5290.

Dumitru, Cristina, Ghaida Muttashar Abdulsahib, Osamah Ibrahim Khalaf und Akram Bennour. 2025. „Integrating Artificial Intelligence in Supporting Students with Disabilities in Higher Education: An Integrative Review." *Technology and Disability*, online first, 27. Juni 2025. https://doi.org/10.1177/10554181251355428.

Friedrich, Julius-David, Jens Tobor und Martin Wan. 2024. *9 Mythen über generative KI in der Hochschulbildung*. HFD-Diskussionspapier 29. Berlin: Hochschulforum Digitalisierung.

Georgiou, Georgios P. 2025. „ChatGPT Produces More ‚Lazy' Thinkers: Evidence of Cognitive Engagement Decline." arXiv:2507.00181.

Hochschule Bayern. 2025. *KI-Leitlinie Hochschullehre. Empfehlungen zum Umgang mit KI in der Lehre der bayerischen Hochschulen für angewandte Wissenschaften und Technischen Hochschulen*. München: Hochschule Bayern e. V.

Wannemacher, Klaus, Elke Bosse, Maren Lübcke und Alena Kaemena. 2025. *Wie KI Studium und Lehre verändert. Anwendungsfelder, Use-Cases und Gelingensbedingungen*. HFD-Arbeitspapier 87. Berlin: Hochschulforum Digitalisierung.

Humboldt, Wilhelm von. 1809/1903. „Über die innere und äußere Organisation der höheren wissenschaftlichen Anstalten in Berlin." In *Wilhelm von Humboldts politische Denkschriften*, Bd. 1: 1802–1810 (= *Gesammelte Schriften*, Bd. 10), herausgegeben von Bruno Gebhardt, 250–260. Berlin: B. Behr.

Kleist, Heinrich von. 1805/1990. „Über die allmähliche Verfertigung der Gedanken beim Reden." In *Sämtliche Werke und Briefe in vier Bänden*, Bd. 3: *Erzählungen, Anekdoten, Gedichte, Schriften*, herausgegeben von Klaus Müller-Salget, 534–540. Frankfurt am Main: Deutscher Klassiker Verlag.

Kosmyna, Nataliya, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein und Pattie Maes. 2025. „Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task.“ arXiv:2506.08872.

Liessmann, Konrad Paul. 2006. *Theorie der Unbildung. Die Irrtümer der Wissensgesellschaft*. Wien: Zsolnay.

Miao, Fengchun und Wayne Holmes. 2023. *Guidance for Generative AI in Education and Research*. Paris: UNESCO.

Münch, Richard. 2009. *Globale Eliten, lokale Autoritäten. Bildung und Wissenschaft unter dem Regime von PISA, McKinsey & Co*. Frankfurt am Main: Suhrkamp.

Nida-Rümelin, Julian. 2014. *Der Akademisierungswahn. Zur Krise beruflicher und akademischer Bildung*. Hamburg: edition Körber-Stiftung.

Risko, Evan F. und Sam J. Gilbert. 2016. „Cognitive Offloading.“ *Trends in Cognitive Sciences* 20 (9): 676–688. https://doi.org/10.1016/j.tics.2016.07.002.

Rosa, Hartmut. 2005. *Beschleunigung. Die Veränderung der Zeitstrukturen in der Moderne*. Frankfurt am Main: Suhrkamp.

Technische Universität München. 2025. „Erstmals mehr als 100 neue Start-ups in einem Jahr an der TUM.“ Pressemitteilung, 28. März 2025. https://www.tum.de/aktuelles/allemeldungen/pressemitteilungen/details/erstmals-mehr-als-100-neue-start-ups-in-einem-jahr.

Wissenschaftsrat. 2022. *Empfehlungen zur Digitalisierung in Lehre und Studium*. Drs. 9848-22. Köln: Wissenschaftsrat.

Teil III

Denkfiguren: Utopie, Analogie, Gedankenexperiment

Die Universität als Soziale Plastik – eine utopische Skizze

Nicolaus Wilder

„wer nicht denken will fliegt raus“

Joseph Beuys, 1977

Der eigentliche Skandal von KI für die Universität besteht nicht darin, dass Maschinen inzwischen Texte schreiben, Argumente simulieren und Informationen in Sekunden aggregieren können, sondern darin, dass unter ihrem Druck sichtbar wird, wie unklar geworden ist, was Universität überhaupt noch sein soll. Wo sie sich nur noch als effiziente Infrastruktur einer vermeintlichen Wissensproduktion versteht, ist KI ein willkommenes, mächtiges Instrument; wo sie aber ein Ort des Urteilens, des Widerspruchs, der Verantwortungszurechnung und der gemeinsamen Erprobung alternativer Möglichkeiten ist, ist sie bloß neues Material. Die Frage lautet deshalb nicht, ob KI die Universität rettet oder zerstört. Die Frage lautet, welche Universität unter ihrer Existenz gewollt, verteidigt und gestaltet werden soll. Dieser Essay schlägt vor, die Universität nicht primär als Organisation, Dienstleisterin oder Innovationsmaschine zu begreifen, sondern als Soziale Plastik: als eine von Menschen gemeinsam geformte soziale und normative Gestalt, in der Wahrheitssuche und Weltgestaltung, Kritik und Kreativität, Freiheit und Verantwortung aufeinander bezogen bleiben.

Eine solche Utopie kann jedoch nur begründet entworfen werden, wenn zuvor deren Möglichkeitsbedingungen in Vergangenheit und Gegenwart rekonstruiert wurden.[1] Sie gewinnt ihre kontrafaktische Begründetheit also aus dem Faktischen, aus dem, was in Erfahrung bereits Gestalt angenommen hat, aus der historisch vorliegenden Sinnstruktur – und ist deswegen nicht bloß Hirngespinst wie einige gegenwärtige Visionen von KI-Zukünften zwischen Apokalypse und Paradies. Hermeneutische Zugriffe betonen, dass Sinn dabei weder fertig noch unverhüllt vorliegt, sondern in geschichtlichen Verständigungsprozessen gebildet, vermittelt und verändert wird. In diesem Sinn ist auch Universität kein bloßer Organisationsbegriff. Er bezeichnet eine soziale und normative Ordnung, eine Praxis und Lebensform, in der Menschen Wissen – im Folgenden verstanden als kritisch reflektierte Meinung oder Information – hervorbringen, prüfen, weitergeben und öffentlich verantworten. Diese Praxis ist in ihrer konkreten Ausgestaltung historisch zwar höchst vielfältig, bleibt aber stets gebunden an einen Kern, ein Wesen, eine Idee der Universität, die invariante Struktur in aller Varianz, von der ausgehend dann alternative Gestaltungsmöglichkeiten erschlossen werden können.

Als ideengeschichtlich heuristischer Fixpunkt dieser Struktur kann Platon dienen. Mit der Akademie als Lebens- und Lerngemeinschaft Gleichberechtigter sowie der Grundlegung

[1] Zur begründeten Erschließung von Möglichkeiten vgl. Kap. 2.1 in Wilder, N. (2022). *Superheld*innen – Gottheiten der Gegenwart? Ein pädagogisch-historischer Versuch zu den Orientierungsmöglichkeiten des Mythischen und Utopischen*. Universitätsverlag Kiel. https://doi.org/10.38072/978-3-928794-85-5

der Bildungsidee im Höhlengleichnis[2] werden zwei Achsen angelegt, die bis heute das Selbstverständnis universitärer Bildung strukturieren: das Streben nach Wahrheit einerseits und das verantwortungsvolle Streben nach Gerechtigkeit, mithin nach einer besseren Ordnung des Zusammenlebens, andererseits. Beide gehören nicht in getrennte Sphären, weil Erkenntnis hier nicht als Sophisterei oder Selbstzweck erscheint, sondern als Bedingung eines besseren Lebens: Wahrheitssuche ist nicht bloß kontemplativ, sondern an die Frage – den höchsten Zweck – gebunden, wie Menschen zusammenleben sollen; und umgekehrt kann Gerechtigkeit nicht ohne Einsicht und Prüfung gedacht werden. An diesem Fixpunkt zeigt sich, dass utopisches Denken der Idee der Universität von Beginn an eingeschrieben ist, was Platons *Politeia* idealtypisch illustriert.

Institutionell entsteht Universität als *universitas magistrorum et scholarium* erst im Hochmittelalter. Gerade dort wird ergänzend sichtbar, dass es um eine korporative Form geht: um Selbstverwaltung, um Curriculumshoheit, um Verfahren der Gradevergabe, um eine Öffentlichkeit des Disputierens. Die Moderne radikalisiert dann insbesondere Autonomie, Lehre und Forschung als zusammengehörige Praktiken.[3] Dabei bleibt aber erhalten, dass Wissenschaft nicht einfach Technik der Informationsproduktion ist, sondern eine Praxis, die an Wahrheit, Kritik und Begründbarkeit gebunden ist, auch in ihrer Auseinandersetzung mit Möglichkeiten. Wahrheitssuche hat gesellschaftliche Konsequenzen, und gesellschaftliche Orientierung braucht Formen des streitbaren Prüfens. Auch hier zeigt sich das utopische Moment der Universität: nicht als Flucht aus der Wirklichkeit, sondern als institutionalisierte Praxis, die das Gegebene als Basis nimmt, um von dort aus alternative Möglichkeiten zu erschließen, insbesondere Möglichkeiten einer besseren, also gerechteren Welt.

Im Kern ist Universität also eine *Gemeinschaft*, die Bildung und Forschung als Urteilsbildung organisiert. Wahrheit und Wissen sind dabei nicht Mittel intellektueller Selbstbefriedigung und -darstellung, sondern Mittel einer Praxis, die Kritik, Verantwortung und Freiheit so institutionell bündelt, dass Gesellschaft als gestaltbar erfahren werden kann. Darin liegt die Orientierungsfunktion von Wissenschaft in einer säkularen Welt.[4]

Diese Dimension wird jedoch schon länger zunehmend verdrängt: Universität erscheint mehr und mehr als Maschine zur möglichst schnellen und effizienten Anhäufung von Daten, Credits, Publikationen und verwertbaren Kompetenzen. U. a. diagnostizierten Adorno eine „seltsame Schrumpfung des utopischen Bewußtseins“[5] und Bieri eine zeitgenössische

[2] 7. Buch in Platon. (2017). *Der Staat*. Reclam.
[3] von Humboldt, W. (1960). Über die innere und äußere Organisation der höheren wissenschaftlichen Anstalten in Berlin [1809 oder 1810]. In W. Weischedel, W. Müller-Lauter & M. Theunissen (Hrsg.), *Idee und Wirklichkeit einer Universität. Dokumente zur Geschichte der Friedrich-Wilhelms-Universität zu Berlin* (S. 193–202). Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110848076-011
[4] Mittelstraß, J. (2019). Bildung in einer Wissensgesellschaft. *heiEDUCATION Journal. Transdisziplinäre Studien zur Lehrerbildung*, 21–36. https://doi.org/10.17885/HEIUP.HEIED.2019.3.23942
[5] Adorno in Bloch, E. (1975). Etwas fehlt… Über die Widersprüche der utopischen Sehnsucht. Ein Gespräch mit Theodor W. Adorno. In R. Traub & H. Wieser (Hrsg.), *Gespräche mit Ernst Bloch* (S. 58–77). Suhrkamp, hier S. 60.

Fantasielosigkeit;[6] beides verweist auf einen Verlust fundamentaler Entwurfs- und Möglichkeitsfähigkeiten, den Universität nicht nur beklagen, sondern praktisch bearbeiten muss, will sie Universität bleiben.

In einer Utopie von Universität wird diese damit selbst zum Ort, an dem Wirklichkeit nicht bloß vermessen, sondern im Medium von Kritik, Möglichkeitsentwurf und Verantwortung überschritten wird – und an dem Kreativität als Gestaltung sozialer Formen wieder zum Kern akademischer Praxis werden kann. Dazu gehören drei Minimalbedingungen.

- *Institutionalisierte Kritik*: Gründe müssen prüfbar sein und Widerspruch muss nicht nur erlaubt, sondern durch Verfahren (Seminar, Disputation, Review etc.) erzeugt werden.
- *Adressierbare Verantwortung*: Autor*innenschaft, Haftung und Revision dürfen nicht im Nebel verschwinden, sondern müssen über Regeln, Zugänge und Korrekturwege zurechenbar bleiben.
- *Autonomie*: Schutz nach außen und Freiheit nach innen, damit Universität nicht zur Kompetenzherstellungsmaschine oder zur Ideologieschule wird. So entsteht ein Raum, in dem Alternativen entworfen, erprobt und verantwortet werden können.

Universitäten selbst definieren sich heute häufig über ihre Outputs: Abschlüsse, Publikationen, Transfer. Doch nichts davon ist spezifisch universitär, solange es als Ergebnis auch anderswo erzeugt werden kann. Ein Zertifizierungsbetrieb kann Abschlüsse vergeben, ein Forschungsinstitut kann publizieren. Universität wird daher erst dann sichtbar, wenn nach der Form gefragt wird, in der solche Ergebnisse zustande kommen. Der entscheidende Punkt ist nicht, *dass* Erkenntnis produziert wird, sondern *wie*: in Verfahren, die Kritik institutionalisieren, Verantwortung adressierbar halten und die Beteiligten nicht als Kundschaft, sondern als Teil einer gemeinsamen Praxis ansprechen.

Damit werden auch Grenzfälle lesbar, die im KI-Diskurs oft als Zukunftsbilder erscheinen. Eine rein KI-getriebene Forschungsplattform kann Erkenntnisleistungen generieren, ohne Universität zu sein, wenn sie weder Lehr-Lern-Beziehungen noch öffentliche Rechtfertigungspflichten integriert. Eine *corporate university* kann Weiterbildung organisieren, ohne Universität zu sein, wenn sie Autonomie zugunsten der Zweckbindung aufgibt. Eine digitale Hochschule jedoch kann durchaus Universität sein, wenn sie u. a. die sozialen Formen des Prüfens, Widersprechens und der Verantwortungszurechnung nicht durch *Services* ersetzt, sondern institutionell organisiert. Nicht das Medium entscheidet, sondern die soziale Grammatik der Praxis. Dass das jedoch eine große Herausforderung darstellt, hat die Coronapandemie in aller Härte offengelegt, denn diese soziale Grammatik ist nicht rein kognitiv-argumentativ. Widerspruch ist auch leiblich-situativ und stets gebunden an die eigene Positionalität und Perspektivität:[7] im Tempo eines Seminars, in Unterbrechbarkeit und Schweigen, in der Erfahrung von Scham und Mut, in Blickregimen und Mikro-

[6] Bieri, P. (2005). *Wie wäre es, gebildet zu sein?* https://www.hoffbauer-stiftung.de/fileadmin/user_upload/hoffbauer/content/bildung/fort_und_weiterbildung/echris/schulwesen-allgemein/Peter-Bieri-wie-wa__re-es-gebildet-zu-sein.pdf

[7] Koerrenz, R. (2023). Der hermeneutische Zirkel. *Pädagogische Rundschau*, *77*(2), 135–163. https://doi.org/10.3726/PR022023.0012

praktiken der Aufmerksamkeit. Wer von Urteilskraft spricht, spricht nicht nur über Argumente, sondern auch über Situationen, in denen Argumente vorgebracht werden.

Digitale Arrangements können solche Situationen stützen (z. B. durch niedrigere Zugangshürden, asynchrone Revision oder neue Formen der Dokumentation). Sie können sie aber gleichermaßen schwächen, wenn Diskurse zu Monologen mit schwarzen Kacheln verkümmern und Kritik in sykophantisches KI-Feedback verwandelt wird. Die Frage ist daher nicht Präsenz versus Digitalität, sondern welche Präsenz- und Leiblichkeitsdimensionen eine konkrete Lern- und Forschungsform benötigt, um Widerstand als Erfahrung zu ermöglichen. Genau daran hängt ein Wesensmerkmal der Universität: dass Kritik nicht nur geäußert werden darf, sondern situativ möglich, zumutbar und adressierbar bleibt. Leiblichkeit stützt hier die soziale Grammatik der Verantwortungszurechnung, weil Sprechen, Widersprechen und Rechtfertigen nicht im luftleeren Raum stattfinden, sondern in riskanten, dichten Situationen.

Diese schemenhafte Rekonstruktion von Universität eröffnet einen Möglichkeitsraum für eine Utopie, in der ihre konstitutiven Strukturmerkmale nicht übersprungen, sondern gestaltet werden. Genau hier wird Joseph Beuys produktiv. Seine berühmte Formel „Jeder Mensch ist ein Künstler“ ist keine ästhetische Provokation, sondern eine Institutionenthese: Kreativität ist nicht das Privileg weniger, sondern ein „existentielle[s] Potential“[8], eine allgemeine menschliche Fähigkeit und Verantwortung, Weltverhältnisse zu formen, konkret: Gesellschaft zu gestalten. Denken, Sprechen und Handeln sind plastische Vorgänge, in denen soziale Form entsteht. Damit verschiebt sich der Maßstab: Nicht die maximale Akkumulation von Informationen und Daten ist das Ziel, sondern die Freisetzung einer Gestaltungskraft, die Wissen und Kritik mittels Kreativität in eine bessere Lebensform aller übersetzt, wie es bereits bei Platon konzipiert war. Kreativ ist hier nicht, wer Varianten produziert, sondern wer aus der Grundkraft der Freiheit heraus soziale Ordnung mitgestaltet und die Folgen dieser Gestaltung öffentlich verantworten kann. In diesem Verständnis ist KI nicht kreativ. Sie kann überraschende Kombinationen erzeugen und Variantenräume erschließen, aber sie erfüllt nicht die Voraussetzung der Freiheit, hat keine Position in und Perspektive auf Welt, aus der heraus überhaupt erst eine neue Perspektive, Verantwortung, Selbstbindung und ein Wollen entstehen können.

Die Chiffre, in der Beuys diesen zutiefst menschlichen Zusammenhang bündelt, lautet *Soziale Plastik*.[9] Sie bezeichnet Gesellschaft selbst als gestaltbares Kunstwerk, als sozialen Organismus, der durch gemeinsame Praxis geformt wird. Praktisch anschlussfähig an eine Utopie der Universität ist das in Beuys’ Idee der *permanenten Konferenz*: Projekte wie die *Honigpumpe am Arbeitsplatz* (documenta 6, 1977) und die Gründung der *Free International University* (FIU) (1973) zeigen, dass die Soziale Plastik nicht nur ein (Ab-)Bild, sondern ein Organisationsmodus ist, um gemeinschaftlich in und an einer freiheitlich-

[8] Beuys zit. n. Harlan, V., Rappmann, R. & Schata, P. (1976). *Soziale Plastik. Materialien zu Joseph Beuys.* Achberg, hier S. 39.
[9] Obraz, M. (2025). *Die Soziale Plastik von Joseph Beuys als Grundlage einer konstruktiven Sozialisation. Die sozialen Aspekte des erweiterten Kunstbegriffes*. Universität Münster. https://doi.org/10.17879/31948438234

demokratischen Gesellschaft zu arbeiten: Gespräch, Streit, Interdisziplinarität und gemeinsames Denken zu gesellschaftsrelevanten Themen ohne Zugangsbeschränkungen werden so selbst zur Produktionsform.

Die analytische Kraft dieser Metapher liegt dort, wo sie in institutionelle Gestaltungsprinzipien übersetzt wird. *(1) Partizipation als Mitgestaltung:* Soziale Plastik ist keine Beteiligung als Zusatz, sondern die organisierte Möglichkeit, dass alle Beteiligten Initiativrechte haben, Konflikte sichtbar machen und Formen verändern können. Dabei geht es nicht bloß darum, die Praxis gemeinsam zu gestalten, sondern auch deren Bedingungen. *(2) Zeitlichkeit und Langform:* Soziale Plastik arbeitet nicht im Takt des Outputs, sondern in Revision, Werkstatt, Wiederholung, d. h. in institutionell geschützter Langsamkeit, in der Irrtum und Korrektur als produktive Kräfte wirken. *(3) Materialität der Mittel*: Infrastrukturen (Räume, Daten, Plattformen, Prüfungsformen) sind nicht neutral; sie schreiben Werte ein. *(4) Konfliktfähigkeit:* Die Form des Gemeinsamen entsteht nicht durch Glättung, sondern durch geregelten Streit.

Damit wird zugleich der mögliche Ort von KI-Systemen in dieser Universitätsutopie sichtbar. Sie können Verfahren ändern, ohne deren Kern zu treffen: Sie können Recherchen beschleunigen, Übersetzungen unterstützen, Varianten sichtbar machen, unliebsame Routinen übernehmen. Solche Verschiebungen sind mit der Universitätsidee kompatibel, solange das Binnenethos der Kritik, die personale Verantwortlichkeit und die soziale Form des Prüfens erhalten bleiben. Problematisch wird KI-Einsatz dort, wo sie als Trägerin institutioneller Fremdlogiken eingesetzt wird. Die Gefahr besteht dann weniger in KI-Systemen als solchen, sondern in der stillen Übersetzung universitärer Zwecke in Steuerungsgrößen. Diese Fremdlogiken lassen sich grob typisieren: Plattformlogik (Lock-in, Abhängigkeit von proprietären Infrastrukturen), Metriklogik (Rankings, Indikatoren, Performance-Taktung), Compliance- und Risikologik (Standardisierung, Absicherungsroutinen) sowie Marktlogik (Kundschaftsmodell und Serviceversprechen).

Ein naheliegender Einwand lautet, dass diese Erosionen älter sind als KI: Bildung wird seit Jahren durch Ökonomisierung, Quantifizierung und Plattformisierung angegriffen, und KI sei nur Symptom. Gerade darin liegt jedoch die Pointe. KI ist nicht der Ursprung dieser Fremdlogiken, aber sie ist ein Katalysator und hoch und schnell skalierbares Material eines ökonomischen Imperialismus, das diese Logiken in den Alltag der Bildungspraxis hinein verlängert, ja geradezu hämmert.

Der Einsatz von KI für die Hochschule der Zukunft ist also ambivalent: Sie kann Autonomie und Urteilskraft unterstützend ermöglichen, aber ebenso Autonomie durch Metriken ersetzen, Urteilskraft standardisieren, Verantwortung entpersonalisieren und die soziale Form des Prüfens in eine bloße Oberfläche verwandeln. Wenn generative Systeme Texte liefern können, wird das Examen nicht dadurch gerettet, dass KI pauschal verboten wird. Es wird dadurch gerettet, dass Prüfungen wieder als Verfahren der Verantwortungszurechnung begriffen werden. Formate wie mündliche Verteidigung, Prozessdokumentation, Quellenkritik, argumentative Gegenrede oder die Fehlersuche an KI-Outputs machen

sichtbar, ob jemand Gründe geben kann. Sie lassen den Einsatz von KI zu, aber verhindern, dass KI die Stelle des Subjekts besetzt.[10]

In der Forschung wird KI dort problematisch, wo Hypothesengenerierung, Textsynthese und Modellplausibilität mit Wahrheit verwechselt werden. KI kann zwar Möglichkeitsräume erweitern (alternative Hypothesen, Simulationen, Varianten), aber sie kann Evidenz, Rechenschaftspflicht und Verantwortung nicht ersetzen. Gerade in den Geisteswissenschaften gilt: Interpretation ist nicht Textaggregation, sondern begründete Sinnsetzung unter Kritikbedingungen. KI darf hier als Irritations- und Variantenmaschine dienen – Sinn bleibt aber stets in der Verantwortung der Person.

In der Governance schließlich ist KI dort heikel, wo sie entscheidungsnah operiert: Zulassung, Ressourcenverteilung, Risikoprofile, Leistungsmetriken. Hier zeigt sich, dass Verantwortungszurechnung nicht nur in Seminaren, sondern in Infrastrukturen und Verfahren hergestellt werden muss. Wer KI-Systeme nutzt, muss wissen, wer sie trainiert, welche Daten zugrunde liegen, wie Entscheidungen überprüfbar und revidierbar bleiben. Die Frage nach proprietären Plattformen ist deshalb keine technische Randfrage, sondern Teil universitärer Autonomie. Gerade hier wird der Beuys'sche Gedanke der Materialität der Mittel institutionell relevant. Eine Universitätsutopie im Sinne der Sozialen Plastik ist daher keine Technikvision, sondern Aufforderung zur institutionellen Selbstgestaltung. KI ist dann nicht die Utopie, sie ist Material, das es so zu gestalten gilt, dass es die Idee der Universität nicht auslöscht, sondern fördert: KI muss so in Lehre, Forschung und Verwaltung eingebettet sein, dass die Strukturen der Verantwortungszurechnung nicht verschwinden, sondern deutlicher werden. Wo Texte entstehen, muss Autorenschaft geklärt bleiben; wo Entscheidungen vorbereitet werden, muss erkennbar bleiben, wer sie trifft, welche Gründe gelten und wie Revision möglich bleibt. Vernachlässigen wir diese Verantwortung und beugen uns dem Druck der Faktizität, z. B. der Ideologie des Akzelerationismus, wird Gespräch zur Oberfläche, Kritik zur Qualitätssicherung, Bildung zur Kompetenzabfrage, Forschung zur Erzeugung anschlussfähiger Ergebnisse. Was in der *Politeia* als Verschränkung von Wahrheit und Gerechtigkeit angelegt war, wird getrennt: Wahrheit wird zur Funktion, Gerechtigkeit zur Kennzahl. Dann wäre die Institution vielleicht noch Hochschule, vielleicht noch Ausbildungsbetrieb, vielleicht noch Forschungsplattform – aber nicht mehr Universität im eigentlichen Sinn.

Die Gegenfigur ist aber keine Rückkehr zur Technikarmut, vielmehr eine Universität, die ihre Konfliktfähigkeit nicht auslagert, sondern organisiert. KI wäre dort nicht dazu da, die soziale Plastik zu glätten, sondern sie produktiv zu machen: indem Zeit für Gespräch freigesetzt, Alternativen sichtbar gemacht, Kritik unterstützt und Verantwortung gesichert werden. So entwirft die Utopie einer Universität als Soziale Plastik eine Zukunft, in der Universität ein zutiefst und genuin menschlicher Ort ist: ein Ort, der Sinn ins Zentrum rückt, an dem verantwortungsvoll und gemeinschaftlich an einer besseren Welt gearbeitet wird, weil Freiheit nicht delegierbar ist und Verantwortung nicht automatisiert werden

[10] Buck, M. F. (2026). *Prompt zurückgewiesen, Prüfung bestanden? Pädagogische Prüfungen in Zeiten generativer KI*. https://doi.org/10.38071/2026-00523-2

kann. KI kann darin ihren Platz haben, aber nicht als Trägerin von Kreativität oder Urteil, sondern als Mittel, das Menschen unterstützt. Darum ist ihr Einsatz strikt bedingt: Er ist nur legitim, solange KI im Dienst menschlicher Freiheit steht – und er endet dort, wo er Streit und Kritik durch Kennzahlen und Standards simuliert; wo Menschlichkeit zugunsten von Effizienz und nutzenmaximierendem Vorteilstausch[11] aufgegeben wird; wo Zukunft der Gegenwart zum Opfer fällt, weil der Takt der Metriken Langform und Revision verdrängt; wo Freiheit einem bequemen Unterwerfen unter politische Vorgaben, Compliance-Routinen und Methodologismen weicht, wo Möglichkeiten reduziert statt erweitert werden.

In der Utopie der Universität als Soziale Plastik bleiben Wissenschaft und Bildung, Wahrheit und Gerechtigkeit, Kreativität und Vernunft, Wirklichkeit und Möglichkeit, Freiheit und Verantwortung stets aufeinander bezogen. KI ist Anlass, um diese Aspekte in ihrer konkreten aktuellen Ausgestaltung neu zu verhandeln, und gestaltbares Material, das man im Dienst dieser Aspekte formen muss, entscheidend für die grundlegende Idee der Universität ist sie jedoch nicht.

[11] Becker, G. S. (2025). *Der ökonomische Ansatz zur Erklärung menschlichen Verhaltens*. Mohr Siebeck. https://doi.org/10.1628/978-3-16-165674-3

Universitätsidee im Spiegel der Sportidee: Eine Analogie als Inspiration

Frank Vohle

1. Über den Tellerrand der Universitätsidee: Analoge Anregungen

Ich gehe mit diesem Essay der Vermutung nach, dass die gegenwärtige Verunsicherung der Universität weniger durch neue Technologien verursacht wird, als durch einen Verlust begrifflicher Klarheit über ihr Selbstverständnis. Die Debatten um Künstliche Intelligenz (KI) in der Hochschulbildung kreisen um Effizienz, Prüfungsformate, Autorschaft oder Täuschung. Doch all diese Fragen setzen voraus, dass klar ist, was Universität – verstanden als Trias von Forschung, Lehre und Bildung – ihrem inneren Anspruch nach eigentlich bedeutet. Genau daran aber scheint es zu fehlen. Die eigentliche Herausforderung des 21. Jahrhunderts liegt womöglich in der erneuten Klärung der „Idee der Universität" selbst.

Um dieser Klärung nicht vorschnell mit normativen Forderungen (z. B. human) oder pragmatischen Maßnahmen (z. B. KI-Führerschein)[1] zu begegnen, kann ein *Umweg über eine Analogie* aufschlussreich sein: der Blick auf ein anderes kulturelles Sinnsystem, das ebenfalls unter massivem technologischem, ökonomischem und gesellschaftlichem Druck steht – der Sport, verstanden als kulturelle Idee. Mit kultureller Idee ist gemeint, einen Unterschied zu machen: zwischen der Idee der Musik und den Musikstücken, die wir z. B. bei Spotify hören, zwischen der Idee der Kunst und den Kunstwerken, die wir uns z. B. im Museum anschauen. Ideen sind nichts Empirisches. Man braucht sie aber als abstrakte Deutungs-, Orientierungs- und Entscheidungsinstanz. Analoges gilt für den Sport. Die kulturtheoretische Deutung des Sports bei Sven Güldenpfennig[2] bietet hierfür ein instruktives Beispiel und einen methodischen Weg zur anspruchsvollen Begriffs- und Grenzarbeit.

2. Sport als Idee: Ein kulturelles Drama unter Bedingungen der Selbstbegrenzung

Güldenpfennig gehört zu den wenigen Sportwissenschaftlern, die den Sport theoretisch (re)konstruieren, statt ihn primär empirisch zu beschreiben. Ihn interessiert weniger, wie vielfältig Sport faktisch betrieben wird. Er fragt, was Sport sein muss, um als Sport im engeren Sinne gelten zu können. Sport erscheint dabei als eigenständiges kulturelles Sinnsystem, vergleichbar mit Kunst oder Wissenschaft oder Recht – mit einem spezifischen Proprium, Eigensinn und eigenen Imperativen, mit teilweise kontraintuitiven ethischen Implikationen.

---

[1] https://marktplatz.uni-jena.de/produkt/genki-fuehrerschein/

[2] Ich kenne Sven Güldenpfennig als Autor aus dem Studium (Güldenpfennig, 1996a/b), habe ihn später am Deutschen Olympischen Institut in Berlin persönlich kennengelernt; ich verfolge seine kulturtheoretischen Studien zum Sport (fast 20 Bücher, vgl. Güldenpfennig, 2008) und einen Austausch mit ihm bis heute. Grundlage dieses Essays ist die „Summa", eine systematische Zusammenstellung seiner zentralen Gedanken (Güldenpfennig, 2023a/b).

Zunächst sind einige methodische Vorüberlegungen zur Analogiebildung anzustellen. Die Analogie ist in der Wissenschaft ein bekanntes und grundlegendes Erkenntnisverfahren. Bei der Analogiebildung wird Wissen aus einem bekannten Quellbereich auf einen noch unbekannten Zielbereich übertragen. Bei dieser Übertragung, dem Mapping, sind *strukturelle und funktionale Ähnlichkeiten* leitend. Die Analogie ist im Prozess der Exploration, Innovation und Kommunikation beliebt, weil sie mit einem Schlag das Unbekannte mit dem Bekannten greifbar macht; sie ist bei der Verifikation geächtet, weil sie nichts beweist und nur Plausibilitäten anbietet. Ein verantwortlicher Umgang mit der Analogie entsteht, indem im Quell- und Zielbereich Strukturähnlichkeiten herausgearbeitet werden und das, was nicht ähnlich ist, also den Unterschied macht, sichtbar wird (Mismatches). Schließlich gilt es, das heuristische Potenzial der Analogie nicht zu überdehnen und den Eigensinn des Zielbereichs nach der analogen Irritation angemessen zu würdigen (Vohle, 2004).

Sven Güldenpfennig hat den Eigensinn des Sports in einem Set von *sechzehn Kriterien* verdichtet, die gemeinsam den semantischen Gehalt der Leitdifferenz „sportlich-unsportlich" bilden. Diese Kriterien sind ausdrücklich als *idealtypische Rekonstruktion* dessen gemeint, was Sport sein muss, *um im engeren Sinne* als Sport gelten zu können. Diesen Sinn braucht man zwingend, wenn man auf der Phänomen-Ebene „Fußballspiel" und „Bauchtanz" beobachtet und entscheiden will, ob diese Beispiele zur Sinnsphäre des Sports gehören oder eben nicht gehören.

Um sich von der Analogie fruchtbar irritieren zu lassen, möchte ich zunächst den Quellbereich (die Idee des Sports) aus Güldenpfennigs aktueller Publikation von 2026 unverkürzt zitieren.

„Sport erzählt dramatische Geschichten, in denen …

(1) zwei oder mehr *Parteien* sich streiten; und zwar in der janusköpfigen Form einer *Kooperenz*, in der die militant anmutende Konkur*renz* des Miteinanderkämpfens überwölbt wird durch die konfliktdämpfende *Koope*ration bei der gemeinsamen Schaffung des Sportwerkes. Dabei ist der letzt-, nicht aber der erstgenannte Aspekt ausschlaggebend für die Sinngebung des Geschehens. Jenes sportliche Duell vollzieht sich

(2) im Medium meist einer auffällig dominanten *körperlichen Bewegung*, wobei der Einsatz des Körpers als *Mittel* der Leistungserbringung die Grenzmarkierung bildet zwischen dem Sport im engen Sinne und sportähnlichen Formen eines vielgestaltigen Sports im weiten Sinne, wo die Einwirkung auf den Körper aus hygienischen, ästhetischen, psychologischen oder sonstigen Gründen *Ziel und Zweck* des Handelns ist; das Duell erfolgt

(3) ohne einen von außen gegebenen realgesellschaftlichen Grund, sondern aufgrund freiwilliger Verabredung zu einem ansonsten nicht notwendigen und anlasslosen Streit; es geht

(4) um das künstliche Streitobjekt Sieg, das primär keinen direkten materiellen Ertrag einbringt, sondern durch Urkunden, Medaillen oder ähnliche Auszeichnungen für die erfolgreichsten Beteiligten lediglich symbolisch und nur im Fall des professionellen Sports sekundär materiell vergütet wird; das Duell wird ausgetragen

(5) unter dem primären Ziel einer *Selbstvervollkommnung* aller Beteiligten mit den Kontrollinstanzen a) Selbstherausforderung, -beobachtung, -anerkennung, b) Wettbewerb und c) Erfolg über die gegnerische(n) Partei(en), die jedoch vollgültigen Anteil an der kooperenten Schaffung des Sportereignisses haben und dabei primär die Rolle der Fremdherausforderung zur vollen Leistungsentfaltung auf der Gegenseite einnehmen; das Duell verkörpert

(6) eine besondere Form von Wett-Kampf, nämlich als *Wette* der Beteiligten, etwas nicht für möglich Gehaltenes dennoch erreichen zu können, und als Versprechen, um diese Wette unter Einsatz aller sportlich zulässigen Mittel kämpfen zu wollen; es unterliegt

(7) einem *spezifischen Umgang mit Grenzen*: a) Anerkennung von dem Menschen generell gesetzten natürlichen Grenzen, b) Setzung von zusätzlichen künstlichen Grenzen und c) volles Ausreizen des durch diese beiden Bedingungen eingegrenzten Handlungsspielraumes

(8) mittels des ausschließlichen Einsatzes aller Beteiligten von selbst praktisch auf dem Platz zu erbringenden *Eigenleistungen*; das Duell steht

(9) unter strikter Regie vereinbarter und von allen Beteiligten anerkannter *Regeln*, wobei kennzeichnend weniger ist, dass Regeln zu befolgen sind, denn dies gilt überall, sondern dass das sportliche Spiel durch mutwillig gesetzte Regeln überhaupt erst *entsteht* und ohne Regelbefolgung *sinnlos* wird – Regelverletzung bedeutet hier nicht nur moralisches oder rechtliches Versagen, sondern Sinnaufhebung des Geschehens insgesamt; man duelliert sich

(10) unter der Herrschaft des durch diese Regeln begründeten Prinzips sportlicher *Fairness*, das – über die ökonomische Bedeutung von Fairness als gerechten Tausch hinausreichend – von drei Geboten bestimmt ist: a) bei der Verfolgung der sportlichen Ziele nur die durch diese Regeln zugelassenen Mittel einzusetzen, b) die strikte Geltung dieser Regeln anzuerkennen; c) alle durch diese Regeln zulässigen sportlichen Mittel tatsächlich „rücksichtslos“ mit vollem Einsatz auszuschöpfen nach dem Motto „Leistungszurückhaltung ist unfair“;

(11) diese Regeln haben *konstitutive* Bedeutung, insofern deren Geltung das spezifisch sportliche Spiel überhaupt erst hervorbringt, während in anderen gesellschaftlichen Feldern ein schon immer stattfindendes (z.B. ökonomisches oder politisches) Handeln durch andere („regulative“) Regeln „nur“ besonders reguliert wird, um es menschengerechter zu gestalten;

(12) hier gilt ein striktes Primat des *Prinzips Sport* (Lebens- und Gesundheitsgefährdung sowie nicht durch Regeln sanktionierte Gewalt bleiben weitestmöglich ausgeschlossen) vor dem *Prinzip Risiko* (bei dem Gefährdungen bewusst eingegangen werden); man tritt an

(13) unter dem maßgeblich handlungsleitenden, aus der englischen Sportidee geerbten Antrieb, *‚to play the game – and to play it well‘*; das Ganze geschieht

(14) mit dem Letztziel einer gemeinsamen Hervorbringung von Sportereignissen als ästhetischen Werken *(‚Sportwerken‘)* durch die beteiligten Gegner*innen: nämlich als das Erzählen von sporteigenen Geschichten, in dem das Bemühen unter den genannten Kriterien kulminiert und mit dem diese Sport- wie andere Kunstwerke unter einem *Primat der Formgestaltung* als maßgeblichem Kriterium der Zugehörigkeit zur Sphäre der Kunst stehen; dies gilt

(15) mit der Einschränkung, dass das sportliche Werk, so wie die Schöpfungen aller performativen Künste, eine *flüchtige Konsistenz* aufweist, also mit dem Abpfiff unwiederbringlich vergangen ist und nicht einmal, wie manche jener engsten Verwandten, reproduzierbare Spuren hinterlässt – ein nachlesbares Drehbuch, einen dramatischen Text, eine musikalische Partitur, eine notierte Choreographie oder ähnliche beständige Korrelate; das Duell setzt

(16) mit den genannten aufwendigen Bedingungen zwar ein hoch komplexes Geschehen in Gang, mutet aber für alle Beteiligten (also für Akteure wie Beobachter) letztlich doch *einfach nachvollziehbar* an. Diese scheinbare Einfachheit des vordergründigen Geschehens legt irreführende Schlussfolgerungen nahe, in denen verkannt wird, wie voraussetzungsreich der dem sportlichen Geschehen zugrundeliegende Hintergrundsinn tatsächlich ist: Dabei wird oft unterstellt, sportliche Ereignisse seien trivial, trügen keinen eigenen Wert in sich und bedürften künstlicher Aufwertung durch Zuschreibung von außersportlichen Werten, um in einer Kulturgesellschaft ernstgenommen werden zu können. Zudem bedürfe es keines intellektuell anspruchsvollen Interpretationsaufwandes, so dass jedermann beliebig drauflosreden und -handeln könne ohne Respekt vor den Anforderungen, die der sportliche Eigensinn trotz alledem stellt, wenn es sportsinn-gerecht zugehen soll.

Das heißt: Die vordergründige Leicht*verständlichkeit* provoziert Leicht*fertigkeit* im Beobachten, Deuten, Verstehen, Schlussfolgern und Urteilen über das, was den Sport und seine allein aus seinem Eigensinn gespeiste gesellschaftliche Bedeutung und die Verantwortung ihm gegenüber ausmacht." (Güldenpfennig, 2026, S. 2-3; vgl. auch Güldenpfennig, 2023a, S. 344f.)

Güldenpfennig charakterisiert den Sport an anderer Stelle (2023b) zugespitzt als „un-ehrlich, un-friedlich, un-gerecht, un-sozial, un-gesund, un-ökonomisch, un-demokratisch, un-politisch, un-nütz, ja un-sinnig". Was wie eine Provokation klingt, ist eine präzise Grenzmarkierung: Der Sport folgt keiner universalistischen Gesellschaftsethik, sondern einer *funktionsgebundenen Binnen- oder Bereichsethik*, die allein darauf zielt, das sportliche Spiel sinngerecht zu ermöglichen. In dieser „kontraintuitiven" Ethik liegt sein spezifischer Beitrag zur Humanität: Frieden durch Un-Frieden!

Gehört nun ‚Bauchtanz' zur Sinnsphäre des Sports? Bereits ein erster Blick auf das Kriterien-Set macht deutlich, dass Güldenpfennig den Sport nicht über äußere Zwecke – Gesundheit, Erziehung, Integration oder Unterhaltung – legitimiert. Sport erscheint vielmehr als hochgradig reguliertes, künstlich eingerichtetes und zugleich kulturell ernst zu nehmendes Sinngefüge, dessen Bedeutung allein aus seinem Eigensinn erwächst.

Charakteristisch ist dabei dreierlei: *erstens* die strikte Zweckfreiheit des sportlichen Geschehens gegenüber realgesellschaftlichen Anforderungen; *zweitens* der Primat der Selbstvervollkommnung, der alle anderen Kontrollinstanzen – Wettbewerb, Erfolg, Anerkennung – hierarchisch unterordnet; *drittens* die konstitutive Rolle selbstgesetzter Regeln und Grenzen, ohne die sportliches Handeln unfair, ja sinnlos würde.

Der Bauchtanz – ganz im Element der körperlichen Bewegung (da liegt die oberflächliche Ähnlichkeit) – verstößt gleich gegen eine ganze Reihe von Sinnkriterien: Die Bewegung ist hier Zweck und nicht Mittel; sie diente früher als Fruchtbarkeitsritual, heute dient sie der Gesundheit oder der Feierlichkeit bei Festen. Der Bauchtanz hat keine Parteien, funktioniert nicht nach „Kooperenz", ist weder Wette noch Kampf.

3. Ertrag für die Idee der Universität

Wenn die zuvor rekonstruierten Kriterien des Sportsinns als Sinnstruktur und nicht als normatives Raster gelesen werden, entsteht ihr Ertrag für die Idee der Universität nicht durch die Übertragung einzelner Bestimmungen. Erkenntnisproduktiv wird die Analogie vielmehr dort, wo sich „sinnkritische Knotenpunkte" zeigen. Den Vergleich von Universitäts- und Sportidee nehme ich im Folgenden in Form von sechs Bedeutungsclustern vor, die ich drei Ebenen zuweise.[3] Diese Ordnung soll helfen, den inneren Anspruch universitärer Bildung im KI-Zeitalter präziser zu bestimmen.

[3] Vgl. Reinmann & Vohle (in Druck). Wir haben in diesem Text das semantische Feld um den Wertebegriff analysiert und auf der Grundlage einer humanistischen Wertebasis drei semantische Cluster bestimmt, die wir Sinnsphäre, Regelsphäre und Handlungssphäre nennen.

3.1 Sinnebene: Idee und Zweck

(1) Eigensinn: Universität ist keine Dienstleistung! Wie der Sport nicht deshalb existiert, um gesund, nützlich oder sozial zu sein, existiert die Universität nicht primär, um verwertbares Wissen, Innovationsoutput oder Kompetenzzertifikate zu liefern. KI verstärkt die Versuchung, Universität als kognitive Dienstleistungsinfrastruktur zu interpretieren. Der (zum Sport) analoge Einspruch lautet: Der Sinn universitärer Bildung liegt im nicht delegierbaren Erkenntnisvollzug – nicht in seiner externen Verwertbarkeit. Ihr Wert ist indirekt, nicht instrumentell.

(2) Selbstvervollkommnung: Bildung transformiert das Subjekt! KI kann Probleme lösen helfen, aber sie verändert nicht das Subjekt im Vollzug des Problemlösens. Bildung durch Wissenschaft zielt jedoch genau darauf: auf die Transformation des Urteilenden und Handelnden. Analog zum sportlichen Leistungsversuch zählt die eigenständig durchlaufene Anstrengung mehr als der Besitz des Resultats. Leitfrage künftiger Lehre muss daher sein: Welche Formate erzwingen noch persönliche Erkenntnisarbeit statt delegierter Produktion von Resultaten?

3.2 Regelebene: Prinzipien, Ethiken, Grenzen

(3) Kooperenz: Bildung ist nicht a-sozial! Sport ist „kooperenzhaft" verfasst und das heißt, Konkurrenz und Kooperation sind nicht zu trennen; der Gegner ist weniger Hindernis als Mitproduzent des Ereignisses. Übertragen auf die Universität bedeutet das: Bildung entsteht im Widerstreit mit anderen. Argumente gewinnen Geltung im öffentlichen Raum, nicht im isolierten Selbstverhältnis. Personalisierte KI-Lernprozesse steigern Effizienz, bergen jedoch die Gefahr der Entkopplung. Wenn Bildung zur individuell optimierten Interaktion mit einem Algorithmus wird, verliert sie ihre intersubjektive Spannung. Die Analogie zum Sport erinnert daran: Bildung ist keine Sololeistung. Sie braucht Differenz, Widerstand, Öffentlichkeit – und reale Gegenüber.

(4) Binnenethik: Praxisregeln vor Erfolgskriterien! Der Sport ist nicht gerecht im gesellschaftlichen Sinn, sondern regelgerecht im Spiel-Sinn. Übertragen heißt das: Universitäre Praxis darf nicht allein nach Fairness-, Gleichheits- oder Effizienzkriterien der Gesellschaft beurteilt werden. Sie muss zusätzlich an epistemischen Binnenregeln gemessen werden wie Eigenleistung, Zurechenbarkeit, methodische Strenge, argumentative Verantwortbarkeit und Wahrheit. Unter KI-Bedingungen verschiebt sich hier der Prüfstein: Neben das Plagiat tritt die Täuschung über akademische Potenz, analog zum Doping.

(5) Grenzsetzung: Schutz des Unersetzbaren! So wie der Sport sich gegen sachfremde Überformung und Zugriffe seines Spielfeldes[4] wehrt, muss die Universität „Spielzonen" sichern, in denen Erkenntnis unter Echtzeit-, Präsenz- und Responsivitätsdruck entsteht.

[4] Dass Donald Trump als US-Präsident bei der Fußball-Weltmeisterschaft 2026 die Aufhebung der Sperre nach einer roten Karte gegen den US-Stürmer Folarin Balogun „forderte", ist ein gutes Beispiel für einen „sachfremden Zugriff". Obwohl ein solcher politischer Zugriff nach den Statuten der FIFA ausgeschlossen ist, sieht die Realität dann doch ganz anders aus. Die internationale Empörung darüber zeigt allerdings, dass die Idee des Sports als Maßstab funktioniert: Erst sie macht den Vorgang als Grenzverletzung erkennbar.

KI-freie oder KI-limitierte Formate haben mit Technikverweigerung nichts zu tun. Sie sind Bedingung der Möglichkeit von eigener, d. h. selbstproduzierter Erkenntnis. Entscheidend für Erkenntnis und Bildung an der Universität ist, dass Denken sichtbar, riskant, dialogisch und situationsgebunden bleiben.

3.3 Handlungsebene: Erfahrung

(6) Erfahrung / Eigenleistung: der nicht simulierbare Kern! KI kann nahezu alles reproduzieren – nur keine (existentielle) Erfahrung. Erfahrung entsteht im Widerstand: an der Schwierigkeit eines Problems, an der Gegenrede, am Scheitern, an der notwendigen Selbstpositionierung. Universitäre Bildung kulminiert daher im situierten, leibgebundenen Erkenntnisereignis und nicht schon im Wissenszuwachs. Eine Hochschullehre mit Zukunft wird sich daran messen lassen müssen, ob sie diese nicht ersetzbare Erfahrungsdimension systematisch organisiert, oder ob sie diese zugunsten reibungsloser KI-Kooperation preisgibt.

4. Fazit und Schluss

Künstliche Intelligenz entscheidet nicht über die Zukunft der Universität. Sie verschärft nur die Frage nach ihrem Sinn. Wo sich Universität als Dienstleistung versteht, wird KI sie ersetzen. Wo sie sich als Praxis nicht delegierbarer Erkenntnis und Erfahrung begreift, wird KI sie herausfordern – aber nicht ablösen. Wie der Sport seine Würde aus selbstgesetzter Begrenzung gewinnt, so bewahrt die Universität ihre Zukunft aus der Treue zu ihrem Eigensinn.

Wir spielen Fußball mit dem Fuß, obwohl die Hand effizienter wäre. Nicht alles, was leichter ist, gehört zum Spiel. Aber jedes Spiel kann sich auch verändern, mit guten Gründen![5]

Literatur

Güldenpfennig, S. (1996a). *Sport: Autonomie und Krise. Soziologie der Texte und Kontexte im Sport*. Sankt Augustin: Academia.

Güldenpfennig, S. (1996b). *Sport: Kunst oder Leben. Sportsoziologie als Kulturwissenschaft*. Sankt Augustin: Academia.

Güldenpfennig, S. (2008). *Die Würde des Sports ist unantastbar. Zur Auseinandersetzung mit den Mythen des Sports*. Sport als Kultur. Studien zum Sinn des Sports 8. Hildesheim: Arete Verlag.

Güldenpfennig, S. (2023a). *Verteidigung des Sports – Trotz alledem und alledem. Zur Selbstbehauptung gegen Gefährdungen seiner Idee von innen und außen.* Sport als Kultur. Studien zum Sinn des Sports 20. Hildesheim: Arete Verlag.

[5] Inspiriert durch einen wortgleichen Auftakt in der Lektüre von Sven Güldenpfennig.

Güldenpfennig, S. (2023b). *Sportphilosophie – Einführung*. Baden-Baden: Nomos.

Güldenpfennig, S. (2026). *Falsch abgebogen. Zur Vertreibung des Sports aus dem Pantheon der Kultur*. Unveröffentlichter Aufsatz.

Reinmann G. & Vohle, F. (in Druck). Akademische Werte im Zeitalter Künstlicher Intelligenz.

In S. Niedermeier & K. Winkler & H. Mandl (Hrsg.), *Pädagogik und Psychologie im digitalen Wandel* (Arbeitstitel).

Vohle, F. (2004). *Analogien für die Kommunikation im Wissensmanagement. Hintergrund, Mehrwert, Training*. Hamburg: Verlag Dr. Kovač.

Über die Idee der Bildung – eine Nachbetrachtung zum Beitrag „Universitätsidee im Spiegel der Sportidee“ von Frank Vohle

Dominikus Herzberg

Sinnfelder sind die Orte, an denen bestimmte Gegenstände oder Tatsachen auf eine bestimmte Art erscheinen. So definiert Gabriel (2018, S. 91) diesen Begriff für den Neuen Realismus, eine philosophische Denkschule, deren prominenter deutschsprachiger Fürsprecher er ist. Mit Gegenständen ist all das gemeint, „worauf man sich mit wahrheitsfähigen Gedanken beziehen kann“ (Gabriel, 2013, S. 231). „Was Sinnfelder jeweils voneinander unterscheidet, ist ihr Sinn, den wir kennen müssen, wenn wir wissen wollen, in welchem Sinnfeld wir uns befinden“ (Gabriel, 2018, S. 113).

In dem Text „Universitätsidee im Spiegel der Sportidee: Eine Analogie als Inspiration“ entfaltet Frank Vohle das Sinnfeld des Sports – er nennt es Sinnsystem und differenziert es als ein kulturelles Sinnsystem. Mit dem Begriff des kulturellen Sinnsystems zeigt er zweierlei an: Es geht ihm um die Strukturen, die das Sinnfeld auszeichnen, da darin das Potenzial der Übertragung auf den Bildungsbereich liegt (denn Bildung ist definitiv kein Sport), und es geht ihm um eine Charakterisierung des Sinns dieses Sinnsystems. Ein kulturelles Sinnsystem ist ein soziales wie auch ein in kulturellen Praktiken verankertes System, das sich in seiner und durch seine Ausübung und Vollführung selbst hervorbringt. Das ist der Eigensinn des Sports, von dem Vohle spricht. Als Leserinnen und Leser seines Textes sind wir also aufgefordert, das Mittel der Analogiebildung inspirativ dafür zu nutzen, die Strukturen und den Sinn des Sinnfeldes Bildung zu ergründen, um damit weniger in Abgrenzung als vielmehr in Klärung zu gehen, was Künstliche Intelligenz (KI) für das kulturelle Sinnsystem Bildung und seinen Eigensinn bedeutet.

Beim Lesen der von Vohle zitierten sechzehn Kriterien, mit denen der referenzierte Sven Güldenpfennig die Idee des Sports als Sinnfeld zugänglich macht, konnte ich es nicht unterdrücken: Ich tauschte Begriffe und Phrasen aus und las zugleich eine Skizze zur Idee der Bildung! Und ich war von der Kraft dieser Analogie beeindruckt. Sie tat, was sie sollte: Sie brachte mich zum Nachdenken. Denn ganz so leicht ist es nicht mit den Ersetzungen.

- Funktioniert es, „Sport“ durch den Begriff „Bildung“ zu ersetzen?
- Spielt Kooperenz in der Bildung nicht ebenfalls eine wichtige Rolle bei der Schaffung eines Bildungswerks? Oder wirkt ein anderer Mechanismus? Ist der Ausgangspunkt ebenfalls der Streit? Oder ist es die Erkenntnislücke? (1. Kriterium)
- Sind Intellekt und Verstand die entscheidenden Mittel der Leistungserbringung in Analogie zum Körpereinsatz? (2. Kriterium)
- Handelt es sich um die freiwillige Verabredung zur Lückensuche oder zum Lückenfüllen in Erkenntnissen? (3. Kriterium)
- Geht es statt um den Sieg um Erkenntnisgewinn? (4. Kriterium)
- usw.

Die Fragenliste lässt sich mühelos Kriterium für Kriterium fortsetzen. Man merkt, wie sich Widersprüche auftun. Jede mir selbst gegebene Antwort auf eine meiner Fragen muss sich in späteren Kriterien als Passung bewähren. Ich spüre die Nähe zum Sinnfeld der Bildung, dennoch liegt hier nichts auf der Hand, eine schnelle Übertragung durch Wortersetzungen funktioniert nicht. Interessant ist auch, was die Kriterien nicht thematisieren: Es geht hier zum Beispiel mit keiner Silbe um das Trainingswesen oder das Ausbildungswesen zum Sport. So lässt sich auch daraus schöpfen: Muss die Idee der Bildung ihr Trainings- und Ausbildungswesen thematisieren?

Die Kriterien von Güldenpfennig sind inspirierend und verführen zu vielen Fragen. Wie ergibt sich nun eine kohärente Idee der Bildung? Im dritten Kapitel „Analoger Ertrag aus der Sportidee“ geht Vohle dazu über, die Sinnstruktur der Kriterien hervorzuheben, die er anhand dreier Ebenen freilegt: der Sinnebene von Sinn und Zweck, der Regelebene von Prinzipien, Ethiken und Grenzen und der Handlungsebene aus Erfahrung. Und damit kommt er zu sehr klaren Aussagen:

(1) „Der Sinn universitärer Bildung liegt im nicht delegierbaren Erkenntnisvollzug [...].“
(2) „Bildung transformiert das Subjekt“, es zählt „die eigenständig durchlaufene Anstrengung“.
(3) „Bildung entsteht im Widerstreit mit anderen. [...] Sie braucht Differenz, Widerstand, Öffentlichkeit – und reale Gegenüber.“
(4) „Universitäre Praxis [...] muss zusätzlich an epistemischen Binnenregeln gemessen werden wie Eigenleistung, Zurechenbarkeit, methodische Strenge, argumentative Verantwortbarkeit, Exzellenz.“
(5) Die Universität muss „‘Spielzonen’ sichern, in denen Erkenntnis unter Echtzeit-, Präsenz- und Responsivitätsdruck entsteht“.
(6) „Universitäre Bildung kulminiert daher nicht im Wissenszuwachs, sondern im situierten, leibgebundenen Erkenntnisereignis.“

Der Herleitungsprozess ist klug gemacht, die Ableitung aus der Idee des Sports erkennbar. Man könnte zwar auch zu etwas anderen Übertragungen kommen, aber das ist nicht so entscheidend. Das dargelegte Verfahren trägt und es hilft zu erkennen, dass sich damit Probleme beim Einsatz von KI im Bildungskontext deutlich ausmachen lassen. Die kann man sich nun selbst herleiten und zudem abgleichen mit dem, was Vohle dazu anmerkt.

Mit seinem Beitrag ist Vohle etwas Besonderes gelungen: Er nutzt die Analogie als Brücke, um von der Idee des Sports einen Bogen zu schlagen zur kulturellen Sinnstruktur der Bildung, die er in Form der obigen Aussagen verdichtet. Auf einem erstaunlich kurzen Herleitungsweg ist ein Manifest (so nenne ich es) zur Sinnstruktur universitärer Bildung entstanden, das – so meine Hoffnung – zur Reibung und zur Diskussion anstiftet. Die Grenzmarkierungen für den Sinn von Bildung sind gesetzt, und sie fordern heraus.

Bei meiner Nachbetrachtung kam mir noch ein Gedanke, angestoßen durch die Lektüre des aktuellen Buchs von Hartmut Rosa, Situation und Konstellation (2026). Denn auch er zieht den Sport für einen Vergleich heran, in seinem Fall den Fußballsport. Rosa beschreibt in

der Einleitung des Buchs, wie videogestützte Schiedsrichterentscheidungen das Spiel und die Handlungslogiken verändern. Die Videoeinsicht (wie auch andere technische Überwachungsmittel) führt zu einem Anwendungszwang von Regeln, der Augenmaß und Urteilsvermögen der Schiedsrichter außer Kraft setzt. Man kann darin die Idee des Sports in seiner Situiertheit verletzt sehen oder auch in seiner Regelkonformität vollendet sehen. Rosa bringt es in der Gegenüberstellung von Situation und Konstellation treffend auf den Punkt. Aber immerhin: Noch sind es die Menschen, die hier Fußball spielen. Das Sinnsystem des Fußballsports ist nicht grundsätzlich infrage gestellt.

Die Begriffe Situation und Konstellation helfen einem auch beim Verständnis von generativen Sprachmodellen, die Texte für jeden Bedarf in allen erdenklichen Stilen und Formen produzieren. Die diesen Sprachmodellen zugrunde liegenden neuronalen Netze erlernen es, aus den unzähligen Konstellationen von – in diesem Fall – Wortfolgen in Texten (ihrem Trainingsmaterial) Muster und Strukturen abzuleiten, die sich zu situativem Wortfolgenwissen destillieren. Das ist genau das, was Sprachmodelle so attraktiv macht: ihr Vermögen, sprachlich situativ (bezogen auf die vorliegenden Wortfolgen) angepasst zu reagieren. Damit können Studierende Texte produzieren, die sprachlich etwas vortäuschen, was sie selbst nicht als Gedanke, Struktur, Aufbau, Argumentation, Formulierung und Sprachgebrauch vollzogen haben. Das widerspricht eindeutig der von Vohle dargelegten Sinnstruktur von Bildung. Die Situiertheit, Nicht-Delegierbarkeit, das Prinzip der Verantwortung usw. von Bildung sind empfindlich betroffen. Es sind nicht mehr Menschen, die sich auf Bildung als Sinnsystem einlassen, sondern von Menschen bediente Maschinen, die Bildung simulieren.

Können diese Textproduktionsmaschinen denn auch im Sinne von Bildung eingesetzt werden? Kann die Textmaschine nicht auch Bildungsprozesse unterstützen, Texte neu und anders zugänglich machen, sie befragbar machen? Ist es nicht begrüßenswert, dass nun der Nachteil der Sprachbarriere beim Publizieren im Englischen gegenüber Muttersprachlern deutlich geringer geworden ist? Lassen sich mit diesen Maschinen nicht wunderbar viele hilfreiche Übungsaufgaben und -materialien erstellen?

Hier kehrt die Leerstelle wieder, das, was Güldenpfennig nicht thematisiert: das Trainingswesen. Es ist vielleicht kein Zufall, dass die bildungsförderlichen Einsatzszenarien offenbar alle dort liegen – beim Üben, bei der Vor- und Nachbereitung, beim Zugänglichmachen – und kaum dort, wo sich der nicht delegierbare Erkenntnisvollzug ereignet. Wer sich Übungen erstellen lässt, kann noch trainieren, wer sie lösen lässt, nicht. Und die Übersetzungshilfe beim Publizieren? Es hängt daran, ob der Ausgangstext meiner ist und ob ich für die Übersetzung einstehen kann.

Aber ist es wirklich so einfach? Die obigen Aussagen zu den Sinnstrukturen von Bildung lassen Probleme erkennen; manchmal muss man genau hinschauen. Es erscheint z.B. durchaus sinnvoll und vielleicht sogar effizient, sich einen Sachverhalt von einem Sprachmodell erklären zu lassen. Allerdings macht das einen Unterschied: Sprachmodelle haben weder Handlungsabsichten noch Urteilsvermögen, noch können sie das zuverlässig textlich simulieren. Man erhält nicht die gleiche Hilfe wie von einem Menschen. Menschen sind

reale Gegenüber, die zu verstehen versuchen, was die Schwierigkeiten bei Verständnisproblemen sind und wie sie helfen können; das ist nicht immer nur eine Antwort, manchmal auch eine Gegenfrage. Vielleicht ist es hier wie beim Fußballspiel: Das Sinnsystem der Bildung ist nicht grundsätzlich infrage gestellt, dennoch ist etwas anders, wenn die Maschine beim „Trainingswesen“ hilft. Den Einsatz von Textproduktionsmaschinen bildungssinnstiftend zu begründen, ist gar nicht so einfach – aber es erscheint notwendig, sich damit intensiv auseinanderzusetzen. Um auf Gabriel zurückzukommen: Was Sinnfelder unterscheidet, ist ihr Sinn, den wir kennen müssen, um zu wissen, wo wir uns befinden. Ein maschinell erzeugter Textbeitrag erscheint im Sinnfeld der Bildung – ob er ihm auch angehört, entscheidet sich daran, ob sich ein Mensch an ihm gebildet hat. Das herauszufinden ist die Herausforderung, zu der mir Vohle etwas Hilfreiches an die Hand gibt.

Literatur

Gabriel, M. (2013). Die Erkenntnis der Welt: Eine Einführung in die Erkenntnistheorie. 4. Aufl., Freiburg, München: Karl Alber

Gabriel, M. (2018). Warum es die Welt nicht gibt. 4. Aufl., Berlin: Ullstein

Rosa, H. (2026). Situation und Konstellation: Vom Verschwinden des Spielraums. Berlin: Suhrkamp

Die körperlose Universität – ein Gedankenexperiment

Gabi Reinmann & Alice Watanabe

Einführung

Was wäre, wenn …? Das Spiel mit Gedanken beziehungsweise eine mentale Simulation dessen, was (noch) nicht real, aber doch möglich und damit denkbar ist, dürfte ein alltägliches Phänomen sein. Systematisch betrieben als Gedankenexperiment kann dies auch in der Wissenschaft zum Tragen kommen (Kühne, 2005). Insbesondere wenn fachlich relevante Gegenstände und Fragestellungen die Zukunft betreffen und multifaktoriell bedingt sind, lassen sich Gedankenexperimente als Instrument theoretischen Forschens einsetzen (Reinmann & Watanabe, in Druck) – auch in Teilen der Bildungswissenschaft. Indem speziell philosophische Gedankenexperimente mit kontrafaktischen Szenarien arbeiten, befreit man sich mit ihnen zudem von „Zwängen der Faktizität" (Niesel, Jelonnek & Wilder, 2025, S. 333). Für eine Untersuchung des Möglichen eröffnet dies einen weiten und fruchtbaren Spielraum. Philosophische Gedankenexperimente in diesem Sinne haben unterschiedliche Funktionen (Bertram, 2019; Pölzler & Paulo, 2021; vgl. Reinmann, 2024): Sie können Phänomene, Prinzipien oder Theorien illustrieren und erklären oder Gründe für oder gegen diese liefern; mit ihnen lässt sich ein Gegenstand gedanklich überprüfen oder kreativ weiterdenken.

Dieser Essay[1] dreht sich um ein Gedankenexperiment zur Zukunft und Idee der Universität in Zeiten von Künstlicher Intelligenz (KI). Wir konzentrieren uns auf das Szenario und geben Hinweise zu dessen Konstruktion. Die dazugehörige Diskussion und Deutung, die wir selbst zu diesem Szenario verfasst haben, lassen wir außen vor; stattdessen möchten wir die Reflexion und im besten Fall einen Diskurs unter Mitgliedern der Universität (in Studium, Lehre, Forschung, Leitung) anregen. Dieser kann unterschiedlich gestaltet sein: als Zwiegespräch mit sich selbst, als zwischenmenschlicher Dialog, oder als Diskussion in größeren Gruppen. Wir haben das Gedankenexperiment mit dem Ziel konstruiert, den Einfluss von Technik, insbesondere von KI, und dessen potenziellen Folgen anhand eines bewusst skurril anmutenden Szenarios zu durchdenken und dabei zu prüfen, wie schlüssig oder akzeptabel sich das Denkbare zeigt und anfühlt. Es geht also nicht darum, eine normativ schon begründete Entwicklungsrichtung zu illustrieren oder davon zu überzeugen.

Die an verschiedenen Stellen des Szenarios platzierten Reflexionsfragen richten sich an keine bestimmte Zielgruppe, sondern verfolgen allein den Zweck, einen Reflexions- und Diskursanstoß zu geben. Im Anschluss an das Szenario geben wir ein paar Hinweise zu dessen Konstruktion. *Unsere* Interpretation entlang einiger aus dem Szenario ausgewählter Themen lässt sich im Beitrag von Reinmann und Watanabe (2026) nachlesen.

[1] Inzwischen ist der Text in der Zeitschrift „Pädagogische Rundschau" publiziert und dort in der vollständigen Fassung zu lesen (Reinmann & Watanabe, 2026), während wir hier im Essay-Band bewusst auf weite Teile unserer eigenen Interpretation des kontrafaktischen Szenarios verzichtet haben, das im Folgenden im Mittelpunkt stehen wird.

Das Szenario

Stellen wir uns vor, wir könnten uns in der akademischen Welt unseres Körpers entledigen und diesen wie eine Hülle abstreifen, um uns ganz dem geistigen Arbeiten zu widmen – ohne Einschränkungen infolge von Biorhythmus, Schmerzen oder körperlichen Unzulänglichkeiten, ohne Verzerrung durch Affekte, Emotionen oder innere Bilder. Stellen wir uns weiter vor, wir könnten in diesem körperlosen Zustand besonders effektiv die Potenziale von KI ausschöpfen, indem wir eine direkte Schnittstelle zwischen menschlichen Gehirnen und Maschinen herstellen. Die Möglichkeiten, Neues zu entdecken und zu erfinden, präzise Analysen durchzuführen, das Wissen zu erweitern und das Können zu perfektionieren, wären gigantisch.

Angenommen, jemand würde eine solche technische Möglichkeit schaffen – exklusiv für Universitäten als dem Ort, an dem die schlauesten Köpfe zu finden und geistige Tätigkeiten bestimmend sind: ein winziger Chip, eingebaut in einem Pflaster, das man sich auf den Arm klebt, und schon löst sich das Gehirn vom Körper und verbindet sich mit allen verfügbaren KI-Systemen. Die Präsentation des Chips an Universitäten löst Staunen, Euphorie und Skepsis aus. Besonders groß aber ist die Neugier – nicht ungewöhnlich, sondern geradezu notwendig für Menschen an einen Ort, an dem Wissen geschaffen und vermittelt wird. Viele würden nur zu gerne den Zustand der Körperlosigkeit ausprobieren und fragen sich:

Wie wäre es, wenn Wissenschaftler rund um die Uhr forschen könnten, ohne sich schlafen legen zu müssen, wenn sich Forschende mit allen Disziplinen verständigen könnten, ohne ihre Präzision einzubüßen, wenn Studierende überall auf der Welt mit anderen lernen könnten, ohne an raum-zeitliche Grenzen zu stoßen, wenn Lehrende allen alles erklären könnten, ohne jemanden benachteiligen zu müssen und so fort?

Die Sache hat nur einen Haken: Ein Testbetrieb, an dem die Mitglieder von Universitäten die Körperlosigkeit in der akademischen Welt zunächst testen und sich dann entscheiden können, ob sie mit dem Chip lehren, lernen und forschen wollen, ist ausgeschlossen. Den Zugang zur Körperlosigkeit gibt es für immer oder gar nicht. Einmal eingeführt, ist der Chip mit der Universität verwachsen und lässt sich nicht mehr entfernen. Allerdings wird den Mitgliedern der Universität ein Gestaltungsspielraum in Aussicht gestellt: Forschende, Lehrende und Studierende haben flexible Zeitkontingente für die Körperlosigkeit. Jede Person kann zwischen 20 und 70 Prozent ihrer akademischen Zeit im körperlosen Zustand in Verbindung mit KI geistig tätig sein. Wer zum Beispiel mit 20 Prozent beginnt, kann einmalig oder dauerhaft den Anteil der Körperlosigkeit bis auf 70 Prozent steigern; dann ist die Grenze erreicht. Umgekehrt kann ein hoher Zeitanteil an Körperlosigkeit auch (wieder) bis auf 20 Prozent gesenkt werden; eine Reduktion darunter ist nicht möglich.

Wie würden Sie entscheiden? Gehen Sie den Weg mit in die Körperlosigkeit, um dabei zu sein, wenn sich die akademische Welt dank innovativer Technik transformiert? Oder entscheiden Sie sich gegen den Chip und verzichten darauf, die Wirkungen des körperlosen Zustands zu erleben? Welche Gründe haben Sie für Ihre Entscheidung?

Nehmen wir an, die Entscheidung fällt **für** den Chip der Körperlosigkeit: Trotz erheblicher Proteste einiger Bedenkenträger siegt nicht nur die Neugier. Auch die Sorge, eine

segensreiche Innovation zu verpassen und der akademischen Welt die überall geforderte Transformation aus Mutlosigkeit und Angst zu verwehren, trägt zur Entscheidung bei. Ein Großteil der Lehrenden wie der Studierenden beginnt mit dem kleinsten Zeitkontingent für die Körperlosigkeit; einige Pioniere starten mit mittlerem oder hohem Anteil – vor allem in der Forschung, um mehr publizieren und Drittmittel einwerben zu können. Nach nur wenigen Monaten sind auch in Studium und Lehre Steigerungen zu verzeichnen. Herkömmliches Lehren und Lernen wird immer mehr als Zumutung und Belastung empfunden: Studierende sehen keinen Sinn mehr darin, Texte zu lesen, wenn sie im körperlosen Zustand die Kerngedanken unmittelbar erfassen können. Lehrenden wird es zu anstrengend, Studierenden etwas zu erklären und Verständnishürden überwinden zu helfen, wenn das ohne Körper schneller und reibungsloser geht. Die Lautstärke und der Geruch in überfüllten Hörsälen werden als unangenehm empfunden und gemieden.

Ein Jahr nach Einführung des Chips schöpft die Mehrheit der Universitätsmitglieder das verfügbare 70 Prozent-Zeitkontingent für die körperlose Identität aus; die Akzeptanz der eingebauten Grenze sinkt. Nur eine Minderheit beschränkt sich auf den erzwungenen Zeitanteil von 20 Prozent. Ein Schwarzmarkt entsteht, auf dem Stundenkontingente der Minderheit zum Kauf angeboten werden und reißenden Absatz finden. Die „Körperlichen", wie die Verweigerer inzwischen genannt werden, beobachten eine Explosion geistiger Artefakte und ein Verkümmern des äußeren Erscheinungsbilds derer, die zu lange im körperlosen Zustand blieben – aber sie sind zu wenige, um etwas daran ändern zu können.

*Wie würden Sie die getroffene Entscheidung **für** den Chip der Körperlosigkeit mit den skizzierten Folgen bewerten? Zu welcher Gruppe würden Sie vermutlich selbst gehören: zur Mehrheit oder zu Minderheit? Wie erklären Sie sich die Entwicklung nach der Entscheidung für die Reise in die Körperlosigkeit?*

Gehen wir noch einmal zurück zur Präsentation des Chips der Körperlosigkeit, zum Moment des Staunens, der Skepsis und Bedenken sowie der unbändigen Neugier.

Was wäre nun, wenn trotz Staunen und Neugier an der Universität das Risikobewusstsein einer ausreichend großen Gruppe von Lehrenden, Studierenden und Forschenden die Oberhand behielte und die Entscheidung **gegen** den Chip der Körperlosigkeit ausfiele? Stellen wir uns vor, wie sich die Universität vom Angebot körperloser akademischer Arbeit zurückzieht: Da der abgelehnte Chip exklusiv für Universitäten entwickelt war, wird es für kurze Zeit ruhig um diese Innovation. Doch die Ruhe währt nicht lange. Bald schon ist ein Chip der Körperlosigkeit auf dem Markt, der auf den Industrie- und Finanzsektor zugeschnitten ist. Dort entschließt man sich ohne Zögern für dessen nicht rückgängig zu machende Einführung. Es wächst sogleich der Bedarf an Fachkräften, die sich im körperlosen Zustand mit unmittelbarer KI-Verbindung routiniert zurechtfinden. Ärger und Unverständnis sind groß, als klar wird, dass die Universitäten diesen Bedarf nicht nur nicht decken, sondern grundsätzlich nicht bedienen können.

Es entstehen neue, von der globalen Wirtschaft finanzierte, körperlose Hochschulen, die den herkömmlichen Universitäten schnell den Rang ablaufen – in der Forschung ebenso wie in der Lehre: Sie ziehen viele internationale Interessenten an; ein nicht unerheblicher

Teil unzufriedener Studierender und Forschender verlässt die Universität, um sich der neuen Hochschulgeneration anzuschließen. Als bekannt wird, dass Universitäten den zunächst exklusiven Zugang zum Chip der Körperlosigkeit abgelehnt haben, ist die Kritik groß: Hätten sich die befürchteten Risiken mit dem gewährten Gestaltungsspielraum nicht beherrschen lassen? Die „Körperlichen", wie die massiv geschrumpften Universitäten nun bezeichnet werden, beobachten, wie sich ein kluger Kopf nach dem anderen in die ökonomisch gestalteten Hochschulen begibt – aber sie sind zu schwach, um etwas daran ändern zu können.

*Wie schätzen Sie die getroffene Entscheidung **gegen** den Chip der Körperlosigkeit mit den skizzierten Folgen ein? Wäre vor diesem Hintergrund eine gegenteilige Entscheidung womöglich besser gewesen oder eher nicht? Warum?*

Wie erklären Sie sich, dass die Universität in beiden Fällen (Entscheidung für und gegen die Körperlosigkeit) der Verlierer zu sein scheint? Wie ließe sich das verhindern?

Zur Konstruktion des Gedankenexperiments

Im Mittelpunkt des Szenarios stehen der Körper und physische Erfahrungen als menschliche Eigenschaften. Für den Ort der kontrafaktischen Darstellung wurde die Universität als traditionelle Institution akademischen Handelns gewählt: Geistige Tätigkeiten haben dort im Vergleich zu anderen gesellschaftlichen Kontexten eine besonders hohe Bedeutung und nehmen viel Raum ein. Zumindest auf den ersten Blick ist dieser Ort besonders geeignet, sich der körperlichen Welt zu entziehen.[2] Das Szenario beginnt mit einem (vermutlich) verführerischen Angebot: Universitäten werden dazu eingeladen, einen Chip zu implementieren, mit dessen Hilfe Lehrende, Forschende und Studierende zu einem festgelegten Tagespensum ihre körperliche Hülle ablegen und rein geistig zusammen mit KI produktiv sein können.

Des Weiteren bedient sich das Gedankenexperiment eines Kunstgriffs: Im Verlauf des Szenarios öffnen sich zwei narrative Pfade, was – beim Mitvollzug des Gedankenexperimentes durch die Rezipienten – eine Entscheidung einfordert und das Gedankenexperiment in zwei verschiedene, gleichsam negative Denkrichtungen münden lässt: Mit der Entscheidung **für** die körperlose Universität werden körperliche und zwischenmenschliche Anteile im akademischen Handeln reduziert und bedroht sowie die Pluralität universitären Lebens minimiert. Mit der Entscheidung **gegen** die körperlose Universität verlieren Universitäten an gesellschaftlicher Bedeutung und werden von anderen Bildungseinrichtungen verdrängt.

Mit dieser Entscheidungskonstellation und ihren Folgen thematisiert das Gedankenexperiment, dass Universitäten zwar eigene, rational begründete Entscheidungen im Umgang mit neuen Technologien wie KI treffen können, ohne aber negative Folgen und Begleitumstände gänzlich vorhersehen oder verhindern zu können. Im Gedankenexperiment entwickelt sich eine Paradoxie bzw. ein Dilemma, in dem rationale Entscheidungen ihre

[2] Man mag an dieser Stelle einwenden, dass Disziplinen wie z.B. Ingenieurswissenschaften auch physisch-handwerklich sind; wie aber virtuelle Labore zeigen, nimmt sie das der Möglichkeit der Körperlosigkeit nicht aus.

orientierende Funktion einbüßen (vgl. Honneth & Sutterlüty, 2022): Sowohl die Befürwortung als auch die Ablehnung der Körperlosigkeit in Verbindung mit KI hat gravierende negative Folgen für die Universität.

Die Idee der körperlosen Universität fungiert im Gedankenexperiment als Analogie: Der über ein Pflaster leicht anzubringende Chip steht analog für den niedrigschwelligen Zugang zu (vor allem) generativen KI-Systemen[3]. Die konstruierte radikale Entscheidungssituation für oder gegen diesen Chip zeichnet analog den Handlungsdruck nach, vor dem Universitäten heute oft stehen: Im Zuge der gegenseitigen Konkurrenz unter anderem im Wettbewerb um Fördermittel sehen sich Universitäten oft genötigt zu entscheiden, auf den „KI-Zug" aufzuspringen oder „abgehängt" zu werden. Während der analoge Bezug zwischen dem Pflaster-Chip für Körperlosigkeit und dem leichten KI-Zugang einen größeren gedanklichen Sprung erfordert, ist das beim Entscheidungsdilemma kaum nötig: Dieses ist auch ohne skurrilen Chip schon vielfach real.

In unserer eigenen Deutung und Diskussion (vgl. Reinmann & Watanabe, 2026) gehen wir noch einmal genauer auf den Aspekt der Körperlichkeit bzw. Körperlosigkeit aus philosophischer Perspektive ein und thematisieren damit einen Unterschied zwischen Mensch und Maschine. Des Weiteren greifen wir exemplarisch zwei Themen auf, die wir bei der Entwicklung des Szenarios im Blick hatten: die Abhängigkeit des Menschen von Maschinen sowie die Pervertierung von akademischen Werten. Diese Themen sind nicht erschöpfend: In der Arbeit mit dem Szenario können weitere offene Fragen, Phänomene, auch Paradoxien und analoge Entwicklungen, entdeckt und zum Gegenstand der Reflexion werden.

Schlussbemerkung

Unser Gedankenexperiment lädt dazu ein, über den Zweck der Universität in Zeiten von KI nachzudenken und bietet dazu zwei verschiedene narrative Pfade im Verlauf des Szenarios an. Diese sind weder Trendexplorationen (wie sie mit sozialwissenschaftlichen Techniken erarbeitet werden) noch normativ wünschenswerte Visionen. Vielmehr ist das gesamte Szenario bewusst kontrafaktisch (wie für philosophische Gedankenexperimente üblich) und gleichermaßen dystopisch. Mit der dystopischen Färbung werden tendenziell negative Konnotationen evoziert; hier ließe sich kritisch einwenden, dass diese Konstruktionsentscheidung (zu) einseitig und daher mindestens zu begründen ist. Der Grund, den wir anführen können, besteht darin, dass hochschulpolitische Weichenstellungen derzeit eher von einem KI-Hype, verbunden mit technokratischen Vorstellungen, bestimmt wird, während potenzielle Risiken von KI für die Hochschulbildung unterbelichtet bleiben (Reinmann, Watanabe, Herzberg & Simon, 2025).

Man könnte aber in der Tat ebenso den umgekehrten Weg gehen und ein Gedankenexperiment konstruieren, das bewusst utopische und damit positiv konnotierte Entwicklungen der Universität mit KI skizziert, um auf *dieser* Basis eine mögliche akademische Welt der

[3] Gehring (2024) hat in einem Gedankenexperiment über eine analoge Beziehung zwischen der Veröffentlichung von ChatGPT und der Einführung von Medikamenten die Verantwortungslosigkeit thematisiert, die sich in der unkontrollierten Allgegenwärtigkeit generativer KI manifestiert (vgl. auch Reinmann, 2024).

Zukunft zu imaginieren. Dies wäre deswegen eine Option, weil es nie die Technik an sich ist, die Unheil anrichtet, wie wir es in unserem fiktiven Szenario skizziert haben. Auch wendet sich KI nicht intentional gegen den Menschen und damit gegen das Humane; vielmehr trägt der Mensch selbst dafür die Verantwortung – nämlich mit seinen Entscheidungen, wie er KI konstruiert und verbreitet, wie er sie nutzt, damit Geld verdient etc. Ein utopisches Szenario könnte zum Beispiel dazu anregen, über die faktische wie prinzipiell denkbare Moralität des Menschen nachzudenken. So ließe sich fragen, was möglich wäre, wenn Menschen kein Interesse an Effizienz und Optimierung hätten, nicht bequem und egoistisch wären und niemals anderen schaden wollten: Wie ließen sich dann die Potenziale von KI an der Universität für Wissenschaft und Bildung nutzen? Doch auch in einem positiv konnotierten utopischen Gedankenexperiment bliebe die Frage bestehen, wie sich die Universität und mit ihr Studierende, Forschende und Lehrende durch eine Technik verändern, die dem Menschen immer ebenbürtiger zu werden scheint. Oder würde man im angerissenen utopischen Szenario eine solche KI gar nicht mehr bauen (wollen)?

Abschließend ist es uns wichtig hervorzuheben, dass wir uns nicht gegen den Einsatz von KI in der Hochschulbildung aussprechen. Indem das Szenario unseres aktuellen Gedankenexperiments deutlich von der Gegenwart abweicht, wie auch vermutlich im Prozess der Rezeption irritiert, macht es vielmehr sichtbar und prüfbar, welche Möglichkeitsräume es in der Entwicklung von Universitäten in Zeiten von KI gibt – auch im Hinblick auf bislang (zu) wenig thematisierte Risiken. Unser Ziel ist es, einen reflektierten Diskurs über die neuen technologischen Optionen wie auch Restriktionen und dazu anzuregen, KI in der Hochschulbildung aus unterschiedlichen Perspektiven zu betrachten, zu analysieren und zu gestalten.

Literatur

Bertram, G. (2019). *Philosophische Gedankenexperimente: Ein Lese- und Studienbuch.* Reclam.

Gehring, P. (2024). Rechtspolitische Bemessung möglicher gesellschaftlicher Gefahren digitaler Technologien? Zwei Gedankenexperimente mit anschließender Erwägung. In G. Schreiber & L. Ohly (Hrsg.), *KI: Text: Diskurse über KI-Textgeneratoren* (S. 355–360). De Gruyter.

Honneth, A., & Sutterlüty, F. (2022). Normative Paradoxien der Gegenwart – eine Forschungsperspektive. In A. Honneth, K.-O. Maiwald, S. Spreck & F. Trautmann (Hrsg.), *Normative Paradoxien: Verkehrungen des gesellschaftlichen Fortschritts* (S. 13–38). Campus.

Kühne, U. (2005). *Die Methode des Gedankenexperiments.* Suhrkamp.

Niesel, D., Jelonnek, S., & Wilder, N. (2025). Gedankenexperimente als Methode pädagogischen Denkens – oder: Über die Notwendigkeit des Möglichen. *Pädagogische Rundschau, 79*(3), 321–337. https://doi.org/10.3726/PR032025.0032

Pölzler, T. & Paulo, N. (2021). Thought experiments and experimental ethics. *Inquiry, 67*(1), 355–383.

Reinmann, G. & Watanabe, A. (2026). Die körperlose Universität – ein Gedankenexperiment. Pädagogische Rundschau, 80(4), 411-422. https://www.ingentaconnect.com/contentone/plg/pr/2026/00000080/00000004/art00005#

Reinmann, G. (2024). Gedankenexperimente als bildungstheoretisches Instrument in der Forschung zu Künstlicher Intelligenz im Hochschulkontext. *Impact Free*, 58. https://gabi-reinmann.de/wp-content/uploads/2024/09/Impact_Free_58.pdf

Reinmann, G., Watanabe, A., Herzberg, D. & Simon, J. (2025). Selbstbestimmtes Handeln mit KI in der Hochschule: Forschungsdefizit und -perspektiven. *Zeitschrift für Hochschulentwicklung*, 20 (SH-KI-1), 33-50

Die Idee der Universität Mitte des 21. Jahrhunderts? Kommentar einer Geistes- und Kulturwissenschaftlerin, [hm]niedergeschrieben am 3. Mai 2038

Petra Gehring

[1] „Universität" bezeichnet im Kern kein hochtrabendes Programm, sondern ein institutionelles Arrangement. Dass dem so ist, haben wir in den Jahren ihres Zerfalls bitter begriffen, so wie der Lauf der Dinge sich Ende der 2030 Jahre beschleunigte. Der Idee nach (man sprach von Leitbildern) basierte die Organisation guter Universitäten auf einer ausbalancierten Vielstimmigkeit, wobei die Selbstverpflichtung der Fächer auf einer (ggf. wechselseitig vermittelbaren) methodischen Rechenschaftslegung beruhte. Dass problemgetriebene Interdisziplinarität vor diesem Hintergrund manchmal produktive Horizonterweiterung erwirkte, nicht selten aber auch zu einer gewissen Unschärfe und zu Qualitätskompromissen führte, war schon damals klar. Dennoch funktionierte noch nach Bologna-Reform und agendagetriebener Drittmittelforschung der Mix. Die Universität optierte bereits in den 1990er Jahren „gegen Institution und für Organisation" (Luhmann 1992, S. 98). Aber zumindest in Gestalt von einigermaßen klar unterscheidbaren „Forschungsformen"[1] hat es ein konstruktives Neben- und Miteinander und vielstimmige Forschung gegeben. Freilich reflektierte Wissenschaft schon damals zwar ausgiebig die Technisierung der Gesellschaft. Nicht jedoch die Technikförmigkeit und damit Technologieabhängigkeit (vgl. Husserl 1952) ihrer selbst.

Zum Verhängnis ist den Universitäten in den Jahren nach der aggressiven Freischaltung der ersten, sofort massenhaft genutzten Versionen Generativer KI (GenKI), dann wohl aber das enge Ineinander von Forschung und Lehre geworden. Niemand sah damals die Möglichkeit, der Welle einer „wilden", privat wie in den Schulen ubiquitären Nutzung kommerzieller GenKI-Dienste Einhalt zu gebieten. Man wusste um die vielen für alles Methodische ruinösen Eigenschaften der Tools: fehlende Reproduzierbarkeit, verführerische Geschwindigkeit bei hohem Korrekturzeitaufwand, hohe Anteile an Irrigem, Biases und anbieterseitig gewollte Schlagseiten, dazu (durch oft schwache „Zusammenfassungen") die Abkürzung der Autopsie (also der Inaugenscheinnahme der Quelle) oder allem voran überhaupt: des Lesens und damit der Ruin von Formulierungssinn, Recherche und der Maßstabsfunktion kanonischer Originale – sowie schließlich: die allseitige Gewöhnung an ein Mittelmaß an Genauigkeit und gutem Willen. Im Jahr 2030 kapitulierte bekanntlich dann auch das Publikationssystem. Die ersten Zeitschriften konnten die Qualitätssicherung nicht mehr leisten und ließen GenKI-Systeme als Ko- und Erstautor:innen zu. Seither verschweigen Wissenschaftsautor:innen zumeist, wie sehr das eigene Arbeitsergebnis von ungeprüften Übernahmen aus kommerziellen Maschinen abhängt.

[1] Der Wissenschaftsrat unterschied, noch vor generativer KI, nämlich als die Simulationsforschung aufkam: Beobachten, Experimentieren, Messen, Simulieren, hermeneutisches Interpretieren, theoretische Analyse, Design/Gestalten (vgl. Wissenschaftsrat 2012, S. 35)

Wie gesagt war es aber der Lehrbetrieb, der auch methodenbewusste Forschende binnen sehr kurzer Zeit zum Einlenken zwang. Aus der Lehre stammte die Forderung, die Wissenschaft müsse „beibringen“, wie man GenKI nutzt und somit sofort natürlich auch einen legitimen Ort dieser Tools im Forschungsprozess ausweisen können – um diesen dann auch zu propagieren. Niemand hatte ein gutes Gefühl. Aber dass ein „Verbot“ (selbst auch nur in der Hochschullehre) einem Ausstieg aus jeglichem Fortschritt gleichkäme, verdunkelte die entscheidenden Handlungschancen der frühen Jahre. Hochschulleitungen gingen das Thema nicht beherzt genug an. Didaktiker:innen pfiffen im Walde. Irgendwie würde das schon gehen.

[2] Inzwischen vergleicht man das erste Jahrzehnt mit GenKI mitunter mit der Ausbreitung synthetischer Drogen oder auch mit den ersten Jahrzehnten des um sich greifenden Christentums: einer Fakten leugnenden, von Gerüchten und unprüfbaren Narrativen getriebenen Sekte, deren Abhängigkeit von dem, was die Autorität predigen ließ, den Bildungsanspruch wie auch das Rechts- und das Wirtschaftssystem des Römischen Reiches brach. Wer heute von der Idee der Universität sprechen will, muss zunächst also einmal eingestehen, dass es die Universität von damals, als Institution, als Leitbild, schlicht nicht mehr gibt.

Die Ingenieur- und Naturwissenschaften und auch die empirische Psychologie haben den textlichen Teil ihrer Arbeit wie auch die Durchführung von Versuchen und Erhebungen automatisiert. Durch die Nutzung weniger großer Sprachmodelle und die Etablierung von zumeist durch Großverlage vermarkteten Agentensystemen, die – ebenfalls automatisiert – in der Art eines Systems zweiter Ordnung die Auskünfte verschiedenen GenKI-Systeme auswerten, erfolgt die Personalisierung der Erarbeitung von Forschungsergebnissen nur noch partiell. In der Hauptsache konkurrieren spezialisierte, EU-finanzierte Forschungseinrichtungen miteinander – zu welchen die Hochschulen nicht mehr zählen.

Die vom Wissenschaftsrat seinerzeit unterschiedenen Forschungsformen haben sich durch die großen Modelle zur Ununterscheidbarkeit vermischt. Die Methodenkultur der Geistes- und Kulturwissenschaften allerdings hielt dem Druck, das Schreiben, Lesen und die Leistungskontrollen auf KI umzustellen, vergleichsweise lange stand. Hier wurde 2031 dann auch (mit der Kreativbranche gemeinsam) das Label „hm“, *human made*, eingeführt und zum Bestandteil von Hochschulabschlüssen gemacht. So vergeben immer noch einige Bildungseinrichtungen die begehrten M.A.^{hm}Abschlüsse. Der Bereich der ^{hm}Zertifizierung ist jedoch schon auf der schulischen Ebene überwiegend privatisiert, so dass einschlägige Hochschulausbildungen oft durch private Schulträger organisiert werden – Organisation schlägt Institution. Unter den Top Ten der ^{hm}Bildungsstätten dominieren überdies solche, die durch Großverlage und KI-Anbieter mitfinanziert werden, die natürlich stets auf der Suche nach wirklich exzeptionell guten Trainingsdaten sind. Immerhin sind es heute allein die Geistes- und Kulturwissenschaften, denen man das Prädikat eines wissenschaftlich zu nennenden Denkens zuschreibt. In den Führungsetagen erfolgreicher Unternehmen trifft man rund 85% ^{hm}qualifiziertes Personal, einschließlich Mathematiker:innen. Das Prestige dieser Fächergruppe ist also gestiegen. In absoluten Zahlen wurden jedoch auch diese Fachkulturen weltweit wie auch im nicht-hochschulischen Bildungsbereich stark geschrumpft.

Privatisierung – dies war laut UNESCO ‚der' globale Bildungstrend im vierten Jahrzehnt des 21. Jahrhunderts. Die, weil sie dem zu wenig entgegensetze, schon früh (Klees 2022) kritisierte Weltbildungsorganisation versuchte, den Entwicklungen mittels „Privatization Ethics" gegenzusteuern, was aber weitgehend folgenlos blieb. Auch in Europa dominieren private Bildungsanbieter – bei breitem Einsatz von KI auch im öffentlichen Bereich – die schulische und die berufliche Sozialisation. Dass populistische Parteien nach den öffentlich-rechtlichen Medien auch die steuerfinanzierten Universitäten in Misskredit brachten, ist bekannt. Vielleicht trug dazu aber auch bei, dass das Wort „Qualität" in Diskursen über Forschung eigentlich keinen Ort hatte. Das Konzept der „Methode" jedenfalls hat Produktlinien nicht genug entgegenzusetzen vermocht, für die Geschwindigkeit zählt (sowie niedrigschwellige Konsumierbarkeit) und Qualität nachrangig ist. Erstaunlich ist, dass auch das in den 2010er Jahren noch so zentrale Programmwort der „Nachhaltigkeit" nicht griff, um das zu beschreiben, was Wissenschaft braucht, um Wissenschaft zu bleiben. Vermutlich muss man bilanzieren, dass bereits Anfang der 2030er Jahre wissenschaftspolitische Diskurse primär die Nachhaltigkeitsprobleme energieintensiver KI-Infrastrukturen (also Technik und Kosten), nicht aber mehr die Nachhaltigkeit von „Wissen" (also dessen Technikform, Produktions- und Distributionsbedingungen) zum Gegenstand hatten.

[3] Ich bin Philosophin, ich gehöre zu den noch im Lesezeitalter sozialisierten Zeitzeug:innen meiner Zunft. Ich kann die Frage ein wenig wie eine Ethnologin stellen: Wie sehen heute denn nun die Bildungsorte für die Philosophie, dieses vielleicht undiszipliniertesten und auch hartnäckigsten aller geisteswissenschaftlichen Fächer (und angesichts einer gerade im nicht privaten [hm]Bereich prekärer Finanzierung) aus?

Der breite Strom von mehr oder weniger Trash-TV-artigen Populäraussagen über kanonische Theorieprobleme verengt sich schlagartig, wenn man über Zugang zu Originalquellen verfügt und die vielen – zum Glück ja vorhandenen – KI-Warnsysteme nutzt, die ungeprüfte und Predatory-Inhalte vorfiltern. Die Kosten für einen entsprechenden ausgestatteten Arbeitsplatz sind allerdings hoch. Ebenso ist die [hm]Welt vielfach gespalten, denn für die Kernausbildung gibt es keine Standards mehr. Zertifikate (wie früher ‚der' Universitätsabschluss) sind notorisch strittig; zwischen den Bildungseinrichtungen – Unternehmenscampi, verbliebene öffentliche Hochschulen, subkulturelle Kleinverlage, Lesecafés, virtuelle und physische Archive – finden Kulturkämpfe statt. Rechtfertigungsbedürftig sind insbesondere Fragen der Lizenzierung. Veröffentliche ich auch „für KI", spende ich also Texte und Theoriekompetenz an Maschinen, oder kämpfe ich für [hm]Autonomie und trage bewusst zum Gedeihen eines [hm]Theoriebildungsuniversums bei?

Mir scheint, als entscheide sich die junge Generation heute in Teilen zu einer Flucht in die Präsenz. Es sind nur noch wenige, die Philosophie studieren (und nicht bloß das auf maschinellem Wege leicht erreichbare Label gleichen Namens mitnehmen). Diese wenigen allerdings wollen es wissen. In Kellern, an deren Eingang die smarten Geräte abgegeben werden müssen, finden Lesungen und Diskussionen sowie kollektive Schreibexperimente statt. Antiquariate machen wieder Geschäfte. Und vor allem entwickeln sich schnell voranschreitend neue Jargons, Kennwörter, Idiome, in denen zumindest zeitweilig eine belesene Avantgarde nicht nur das öffentliche Grundrauschen ironisiert, sondern auch genuine

Dinge diskutiert. Und dies rückgekoppelt mit der Fülle einer Überlieferung, die man inzwischen mit voller Absicht den ausgebluteten Einrichtungen der öffentlichen Hand (und damit den Scraping-Bots) entzieht.

Leider steigen gerade im analogen Bereich hierdurch die Knappheiten und die Kosten. Es schwinden die öffentlichen Mittel für physische Archivierung und Bibliotheken, digitale Bibliotheksbesuche bei privaten Anbietern sind teuer – und die letzte Sicherheit der Autopsie von Texten, Bildern oder Tönen fehlt. Es gibt inzwischen Netzwerke älterer Wissenschaftler:innen, die ihre Privatbibliotheken (Print, aber auch alte PDFs und überhaupt Ausdrucke) testamentarisch einem Treuhänder überlassen, der Arbeitsmaterialien für Studierende aller Altersgruppen sammelt. Ich selbst trat vorgestern, am Tag der Arbeit des Jahres 2038, einem solchen Netzwerk zur Vererbung analoger, nicht gemeinfreier, aber geteilter Arbeitsmittel („von-Philosoph:innen-für-Philosoph:innen“) bei.

Literatur

Husserl, Edmund (1952): Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie. In: Gesammelte Werke VI, Den Haag (Nijhoff) ^{2}1975 [Hua VI].

Klees, Steven J. (2022): The Privatization of Education: Reflections on UNESCO’s GEM Report on Non-State Actors. In: Global Education Centre of the Geneva Graduate Institute: NORRAG-Blog. https://www.norrageducation.org/the-privatization-of-education-reflections-on-unescos-gem-report-on-non-state-actors/ [3.5.2026]

Luhmann, Niklas (1992): Die Universität als organisierte Institution. In: Universität als Milieu (hrsg. v. André Kieserling), Bielefeld: Haux, S. 90-99.

Wissenschaftsrat (2012): Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020, Köln: WR. https://www.wissenschaftsrat.de/download/archiv/2359-12.pdf?__blob=publicationFile&v=2 [3.5.2026]

Ausklang

Das Symposium als Bricolage: Eine zu stärkende Praxis für die Zukunft von Universitäten

Alinde Keller

Im Symposium – und im vorliegenden Essayband – fiel mir die Vielstimmigkeit der Beiträge auf. Neben theoretischen Rahmungen und empirischen Erkenntnissen aus heterogenen Disziplinen begegneten mir ungewöhnliche Zeitmodalitäten, z. B. Utopien und Gedankenexperimente, veranschaulichende Zugänge wie eine Analogie oder auch praktische Impulse, z. B. Hinweise zu technischen Eigenschaften generativer KI oder didaktische Ideen. Diese Vielfalt erinnerte mich an die Metapher der *Bricolage* (Lévi-Strauss 1973), die in der Organisationstheorie und -forschung aufgegriffen und weiterentwickelt wurde.

Das Bild der Bricolage zeichnet eine Rationalitätsvorstellung von Organisationen, welche deren Entwicklung weder im Lichte einer maschinellen Logik noch einer vollkommenen Beliebigkeit erscheinen lässt. Bricolage wird als Metapher für eine Praxis genutzt, in der Organisationsmitglieder sich einen Reim auf etwas machen, Entscheidungen treffen und dabei Strukturen ihrer Organisation hervorbringen bzw. in Kraft setzen (z. B. Weick 1993, Duymedjian & Rüling 2012). Sie werden dabei als ‚geschickt Herumprobierende' mit begrenzten Mitteln charakterisiert.

Mit dem Bild der Bricolage möchte ich in diesem Essay beleuchten, inwiefern ich ein Wechselspiel zwischen dem Erproben generativer KI in der Praxis und einer bestimmten Form des *Innehaltens* gerade an Universitäten – die sich auch als Organisationen betrachten lassen – für eine zu stärkende Praxis halte. Dabei meine ich eine Form des Innehaltens, in der sich Akteure mit verschiedenen Sichtweisen begegnen, die Verhältnisse zwischen Problemen und Lösungen hinterfragen sowie diese *diskurssensibel* reflektieren und beurteilen – so wie es auch im Symposium erkennbar war. Diese Praxis erscheint mir aus meiner bildungswissenschaftlichen und organisationspädagogischen Verortung heraus nicht nur für universitätsinterne Entscheidungsprozesse, z. B. zum Festlegen von neuen Regeln zum Umgang mit KI, relevant. Wenn man von einem Auftrag von Universitäten ausgeht, verantwortungsvolle und mündige Gesellschaftsmitglieder auszubilden, und zudem annimmt, dass die meisten Studierenden später in Organisationen tätig sein werden – sei es in der Wirtschaft, im Gesundheitssektor, in der sozialen Arbeit oder im Bildungssystem – dann könnte eine Rationalitätsvorstellung, die mit einer diskurssensiblen Bricolage einhergeht, etwas sein, was sie gerade im späteren Beruf als *selbstbestimmt gestaltende Akteure* benötigen.

1. Denken im Modus der Bricolage: Wildes, aber nicht beliebiges Denken

Denken im Modus der Bricolage mutet etwas wild an. Der Ethnologe und Anthropologe Claude Lévi-Strauss nutzt die Metapher in seinem Buch „Das wilde Denken" (1973) zur Veranschaulichung einer Art wissenschaftlichen Denkens, die er auch eine Strategie

„intellektueller Bastelei“ (ebd., S. 29) nennt. Diese vergleicht er mit dem Handeln technischer Bastler*innen, die mit begrenzten Mitteln herumprobieren und improvisieren. Das französische Wort *bricolage* meint dabei gerade nicht, über alle irgend möglichen Mittel zur plangemäßen Zielerreichung zu verfügen, sondern die Ressourcen und Fähigkeiten zu nutzen, welche gerade zur Hand sind – und sei es auf eine Weise, die Fachexpert*innen abwegig wäre (vgl. ebd.).

Das Verb *bricoler* meint zudem eine Art zu Handeln, bei der immer wieder auf nicht vorhersehbare Ereignisse reagiert wird (vgl. ebd., S. 29). Bricoleure rechnen daher mit Überraschungen und lassen sich von einer eher vagen Zielvorstellung bzw. einem Gefühl für Stimmigkeit und Potenzial des Gesamtwerkes, welches sie schaffen wollen, leiten. Ihr Handeln ist stetige Reorganisation: Sie halten immer wieder inne und prüfen, was gerade entstanden ist (vgl. ebd., S. 33). Dabei interpretieren sie fantasievoll, wofür die vorhandenen Mittel nun gebraucht werden könnten und was dann möglich wäre. Bricoleure fragen z. B.: „Kann ich das gebrauchen? Wie kann ich es einbauen? Bricht dann nicht die ganze Konstruktion zusammen?“ (Ortmann 1995, S. 192). Auch nutzen sie etwa Analogien und Vergleiche (vgl. Levi-Strauss 1973, S. 34). Das Ergebnis wird daher maßgeblich von den zur Verfügung stehenden Mitteln und der Kreativität der Bastelnden geprägt, diese Mittel zu nutzen (vgl. ebd., S. 35). Die Metapher der Bricolage erinnert somit auch an Donald Schöns (1992) Idee von Gestaltungspraktiken als reflektierende Konversation mit den Materialien einer Situation.

Keinesfalls aber handeln Bricoleure völlig beliebig, weil sie die Eigenschaften und früheren Verwendungszwecke ihrer Mittel gründlich kennen und berücksichtigen (vgl. ebd., S. 32). Geschickte Bricoleure zeichnen sich dadurch aus, dass sie ihre Mittel im aktuellen Gebrauch sorgfältig beobachten, den Erfahrungen anderer aufmerksam zuhören und ihr Bauchgefühl sowie ihre praktische Urteilskraft bzw. ‚phronesis‘ – praktische Weisheit oder auch erfahrungsangereicherte Erwartungen – einbeziehen (vgl. Weick 1993, S. 353). Zudem kennzeichnen sich geschickte Bricoleure durch die postheroische wie auch zuversichtliche Haltung „that any enacted structure can be self-correcting if one’s ego is not invested too heavily in it“ (ebd.).

2. Bricolage in Organisationen: Geschicktes Herumprobieren statt planvollen Handelns

Organisationen verstehe ich als spezifische soziale Gebilde, deren Strukturen durch verbindlich gemachte Entscheidungen eine hohe Verlässlichkeit kooperativen Handelns ermöglichen. Dieses kooperative, verbindliche Handeln macht es letztendlich möglich, dass Menschen auf den Mond fliegen, Patienten rund um die Uhr medizinisch versorgt oder komplexe Technologien wie generative KI entwickelt und vertrieben werden – Leistungen, die das Vermögen einzelner Menschen bei Weitem übersteigen.

Gerade Wirtschaftsorganisationen mit technischen Produkten sind traditionell von Hierarchien, kleinteiliger Planung und Kontrolle geprägt. Ihr Erfolg hat historisch gesehen wesentlich zum Einzug strenger ziel- und zahlenbasierter Managementansätze in Organisa-

tionen auch jenseits der Wirtschaft, wie etwa Universitäten, beigetragen. Das Bild der Bricolage entwirft jedoch eine andere Rationalitätsvorstellung von Organisationen. Sie erscheinen darin nicht als planbare und wohlgeordnete Gebilde, die sich an klaren Zielvorgaben ausrichten. Bricolage rückt vielmehr das dynamische Zusammenspiel von etablierten Strukturen und Praktiken des geschickten Herumprobierens mit begrenzten Mitteln in den Blick. Als Bricoleure können sowohl Führungskräfte wie auch Mitarbeitende jenseits hierarchisch festgelegter Befugnisse in den Blick genommen werden, die Einfluss auf Entscheidungen nehmen. Mit ihren Entscheidungen bringen sie die dauerhafte Struktur der Organisation mit hervor. Diese Struktur grenzt den Raum für das tagtägliche Handeln in Organisationen und für weitere (Struktur-)Entscheidungen ein, legt diese aber nicht fest. Zu den Mitteln, die diesen Akteuren zur Verfügung stehen, zählen z. B. ihnen *bekannte* Management-Konzepte, Zahlenwerke, Regeln, Qualifikationen von Mitarbeitenden oder digitale Tools sowie Ideen, Argumente, Ursache-Wirkungs-Erzählungen oder kollektiv geteilte Deutungen. *Begrenzt* sind diese Mittel u. a. deshalb, weil die Akteure sich nicht uneingeschränkt Zeit lassen können und ihre Aufmerksamkeitsspanne limitiert ist. Organisationsmitglieder befragen daher die ihnen bekannten Mittel pragmatisch, was diese in ihrer Situation gerade bedeuten könnten (vgl. Ortmann 1995, S. 191), bevor, während und nachdem sie sich z. B. für Regelungen entscheiden, Softwareprodukte einführen oder Personal einstellen. Mehr noch als technische Bastler*innen allerdings müssen Entscheider*innen in Organisationen mit überraschenden Reaktionen bzw. Intentionen anderer Akteure rechnen; ihr Handeln ist zudem durch Machtgefüge bedingt: „Andere basteln mit und haben vielleicht ganz andere Konstruktionen im Sinn. Die Einsamkeit des [technischen] Bastlers ist dem Manager nicht gegeben“ (Ortmann 1995, S. 193).

Im Bild der Bricolage erscheinen Gestaltungsprozesse in Organisationen folglich weder als völlig beliebig (z. B. Cohen et al. 1988) noch wird ein lineares Input-Output-Verhältnis suggeriert (vgl. Ortmann 1995, S. 189). In den Fokus rückt vielmehr, dass beständige organisationale Strukturen, z. B. Regeln, durch etwas wildes, aber nicht beliebiges Herumprobieren entstehen. Dabei werden sie immer wieder neu, ggf. in leicht variierender Form, zur Geltung gebracht.

3. Das ‚geschickte Herumprobieren‘ mit KI in Universitäten einüben

Die Integration generativer KI in Organisationen aller Art lässt sich plausibel als Bricolage betrachten: KI ist plötzlich da, stellt bisherige Prozesse in ein gänzlich neues Licht und es liegen triftige Gründe vor, sie schnell zu integrieren, auch wenn Zielvorstellungen noch vage sind. Auch in Universitäten gibt es jenseits der Autonomie wissenschaftlichen Personals einen organisationalen Rahmen, in dem kollektiv verbindliche Entscheidungen getroffen werden. Während bereits Prüfungsformate angepasst, Software beschafft oder neue Fortbildungen eingekauft werden, wird gleichzeitig diskutiert, welche Implikationen mit KI einhergehen: Was bedeutet generative KI z. B. für die Formulierungen von Lernzielen, für Curricula oder didaktische Konzepte? Für Evaluationen oder die Leistungsdarstellung von Universitäten? Gar für künftige Finanzierungsmodelle? Welche Bedeutung hatten wiederum diese möglichen Anwendungsfelder in der Vergangenheit? Welchen Zweck

erfüllten z. B. Prüfungsformate, Curricula oder Lernziele bisher – in einer idealen Vorstellung, aber auch in der gelebten Praxis, die das idealistische Ziel häufig verfehlt? Was kann man davon jetzt noch gebrauchen? Wie? Lassen sich neue Potenziale erahnen? Welche? Was möchten wir nun? Was halten wir für wertvoll? Gerade Universitäten haben den Vorteil, über die Ressourcen zu verfügen, mit denen sie sich diesen Fragen während des ‚Herumprobierens' sowohl mit praktischer Klugheit als auch mit wissenschaftlicher Systematik nähern könnten.

Auch im Symposium fragten wir: Für welches Problem ist KI an Universitäten eigentlich eine Lösung? Welche (gesellschaftlichen) Funktionen erfüllen Universitäten? Was davon ist wünschenswert und gewollt – und zwar nicht nur von ‚außen', also von Politik oder Wirtschaft, sondern von uns, dem universitären Personal, selbst? In dieser Art nach dem Verhältnis von Problem und Lösung zu fragen, entspricht einem an Niklas Luhmanns Systemtheorie angelehnten Denkstil. In seinem Essay „Wozu Universitäten?" formuliert der Soziologe und Systemtheoretiker Armin Nassehi in einem solchen Stil folgende Überlegungen für einen universitären (Aus-)Bildungsauftrag:

„Vielleicht kommt es immer weniger auf Bestände als auf die Rekombination von Möglichkeiten an. Vielleicht geht es in der universitären Ausbildung eher um Reflexionsanlässe für die Konstellation von Problem und Lösung statt um fertige Lösungskonzepte. Vielleicht geht es viel stärker um engagierte Distanzierung statt um distanzloses Engagement." (Nassehi 2018, S. 131 f.).

Eine Funktion der von ihm vorgeschlagenen Erarbeitung und Reflexion von Problem-Lösung-Konstellations*varianten* könnte darin liegen, damit anders als vorher urteilen zu können, was wir für erstrebenswerte, stimmige Konstellationen halten. Vielleicht sollten wir als Lehrende und Forschende anlässlich des plötzlichen Auftauchens generativer KI genau das mit Kolleg*innen und vor allem mit Studierenden verstärkt gemeinsam einüben: sich bewusst einen Reim auf etwas machen (hier: auf KI), mutig ausprobierend und dabei immer wieder neu innehaltend, um wahrzunehmen, was gerade entsteht und um zu reflektieren, in welche Richtung es weiter gehen kann. Akteure jeglicher Organisationen – und damit auch Studierende in ihren späteren beruflichen Kontexten – benötigen *auch* eine solche Rationalitätsvorstellung, die mit nicht vorhersehbaren Ereignissen rechnet und welche den Wert eines Innehaltens kennt, in dem Gestaltungsräume neu erkundet werden.

4. Plädoyer für ein diskurssensibles Herumprobieren mit generativer KI

Mit einer *diskurstheoretischen* Brille auf den Umgang mit Wissen in organisationalen Gestaltungspraktiken (Weber 2025) möchte ich ergänzend mit Foucault (1991) fragen: *Welches* Wissen und welche Wissenstypen tauchen *auf welche Weise* im ‚Bastelkasten' bzw. in der Aufmerksamkeitsordnung von Entscheider*innen in Universitäten auf und stehen damit als Mittel zum Basteln bzw. Herumprobieren zur Verfügung? Mit Entscheider*innen meine ich z. B. Lehrende, die etwa Entscheidungen zur Gestaltung ihrer Lehre treffen, aber auch Personen in universitären Gremien, in Stabsstellen oder im Management. Auch

Forscher*innen sind hier gemeint, die etwa entscheiden, welche Forschungsfragen sie wie bearbeiten oder wo und wie sie publizieren. So beobachtete auch Knorr (1979) Formen des ‚tinkering', also des Tüftelns, Bastelns oder Probierens, in der Forschungspraxis.

Im Sinne Foucaults könnte man bei anstehenden Entscheidungen zur Integration von KI fragen:

- Finden vorab Disziplinen mit unterschiedlichen Wissenschaftsverständnissen Gehör? Werden etwa sowohl technische als auch sozialwissenschaftliche oder philosophische Perspektiven berücksichtigt?
- Kommen diese trotz verschiedener Fachhabitus und Denkstile in einen Austausch miteinander?
- Welche Zeitmodalitäten fließen ein? Finden sich neben Rekonstruktionen vergangener Praxis auch erfahrungs- und theorieangereicherte Erwartungen an die „gegenwärtige Zukunft" (Luhmann 2005, S. 112), z. B. in Form von Zukunftsszenarien, Gedankenexperimenten, Utopien oder Dystopien? Wie wird die Differenz zwischen diesen und den begrenzten Möglichkeiten „zukünftiger Gegenwart" (Luhmann 2005, S. 112) verhandelt?
- Welche Rationalitäts- bzw. Nützlichkeitserzählungen tauchen auf?
- Finden sich auch bildbasierte Zugänge? (vgl. z. B. Weber & Heidelmann 2023)
- Sind Überlegungen zur Integration von KI eher von Blickordnungen der Repräsentation, Disziplinierung und Normalisierung oder der Imagination geprägt (Weber 2025)?

Ein in dieser Weise geschärftes Wahrnehmen, warum Probleme als solche gerahmt werden und wie es zu Zielvorstellungen kommt, könnte man *diskurssensible Bricolage* nennen. Diskurssensible Bricoleure nutzen Wissen, Deutungsansätze und weitere Ressourcen fantasievoll, aber keineswegs beliebig, da sie deren vorherigen Gebrauch und ihre Ursprünge kennen, einordnen und kritisch hinterfragen können. Sie bringen wünschbare Zukünfte auch in und durch scheinbar so träge Strukturen wie Organisationen selbstbestimmt mit hervor, und erschiene ihnen ihr Einwirken darauf auch nur marginal.

Auch wenn sich die Praxis der Bricolage teils unbewusst ereignet und man zudem auf der Seite überbordender Reflexivität ‚vom Pferd fallen' kann, weil diese lähmt: Ein begrenztes Innehalten und sorgfältiges Auswählen, wer der Vorbereitung oder Bewertung von Entscheidungen zu Regelungen für generative KI mit welcher Sprechposition mit am Tisch sitzt, können dazu beitragen, gemeinsam Weichen anders zu stellen und am Ende eine andere (organisationale) Praxis entstehen zu lassen.

5. Inwiefern dieser Essayband zum Herumprobieren mit generativer KI einlädt…

Im Arrangement der Essays in diesem Band und in den Beiträgen im Symposium entdecke ich *in a nutshell*: Sie laden zum geschickten Herumprobieren mit Deutungen für den Einsatz generativer KI in universitärer Praxis ein.

Wie etwa von Christian Leineweber in seinem Beitrag vorgeschlagen, könnte die Reflexion bereits entstehender studentischer Praktiken mit generativer KI als Anlass genommen

werden, diese gemeinsam mehrperspektivisch zu deuten und dabei z. B. bildungstheoretisch über die Bedeutung von Bestimmtheit und Unbestimmtheit nachzudenken. Ebenso lassen sich Beiträge mit Zukunftsentwürfen einbinden: Etwa das Gedankenexperiment von Gabi Reinmann und Alice Watanabe, der aus der Perspektive des Jahres 2038 verfasste Kommentar von Petra Gehring oder die utopische Skizze von Nicolaus Wilder. Als wiederum anderer, eher bildbasierter Wissenstyp könnte die Analogie zwischen Universitäts- und Sportidee von Frank Vohle herangezogen werden. Dabei könnten wünschenswerte und dysfunktionale Praktiken diskutiert werden, wozu z. B. die Beiträge von Isabel Steinhardt, Christian Schneijderberg & Oliver Wieczorek oder von Judith Simon Ansporn anregen, um die Urteilskraft zu schärfen.

Gerade im wechselseitigen Kontrastieren, Relationieren und Etikettieren der unterschiedlichen Wissenstypen liegt ein bildungsrelevantes Potenzial (vgl. Dewe et al. 1992). Solche multiperspektivischen Praktiken sind in vielfältiger Weise in Universitäten bereits vorhanden. Es wäre ein Anliegen, diesen nachzugehen und sie sowohl in organisationalen Aushandlungen als auch in der Lehre in diskurssensibler Weise zu verstärken.

Was generative KI selbst in der Praxis der organisationalen Entscheidungsfindung ändert, da sie vermutlich die Aufmerksamkeit von Entscheider*innen weitreichend beeinflusst sowie eigenständig Entscheidungen trifft, und damit mehr als ‚nur' ein weiteres Software-Tool neben vielen anderen ist, dazu vermögen andere fundiertere Aussagen treffen, als ich es derzeit kann. Mein Interesse galt in diesem Beitrag einem Vorschlag, wie man die Beobachtung, dass generative KI universitäre Praxis in neuer Weise mit hervorbringt, deuten und daran gestaltend ansetzen kann.

Literatur

Cohen, M.D., March, J.G. & Olsen, J.P. (1988): Ein Papierkorb-Modell für organisatorisches Wahlverhalten. In March, J.G. (Hrsg.): Organisation und Entscheidung. Gabler.

Dewe, B., Ferchhoff, W. & Radtke, F.-O. (1992). Das "Professionswissen" von Pädagogen. Ein wissenstheoretischer Rekonstruktionsversuch. In B. Dewe, W. Ferchhoff & F.-O. Radtke (Hrsg.), Erziehen als Profession: Zur Logik professionellen Handelns in pädagogischen Feldern. (S. 70–91). VS Verlag für Sozialwissenschaften.

Duymedjian, R. & Rüling, C.-C. (2010). Towards a Foundation of Bricolage in Organization and Management Theory. Organization Studies, 31(2), 133–151. https://doi.org/10.1177/0170840609347051

Foucault, M. (1991). Die Ordnung des Diskurses. Fischer.

Knorr, K. D. (1979). Tinkering toward success: Prelude to a theory of scientific practice. Theory and Society, 8(3), 347–376.

Lévi-Strauss, C. (1973). Das wilde Denken. Suhrkamp.

Luhmann, N. (2005). Weltzeit und Systemgeschichte: Über Beziehungen zwischen Zeithorizonten und sozialen Strukturen gesellschaftlicher Systeme. In N. Luhmann (Hrsg.), Soziologische Aufklärung 2, 5. Aufl. (S. 128–166). VS Verlag für Sozialwissenschaften.

Nassehi, A. (2018). Wozu Universitäten? Eine Legitimationsgeschichte. In A. Nassehi & P. Felixberger (Hrsg.), Kursbuch: 301 Gramm Bildung (S. 115–133). Murmann Publishers GmbH.

Ortmann, G. (1995). Formen der Produktion: Organisation und Rekursivität. Westdeutscher Verlag.

Schön, D. A. (1992). Designing as reflective conversation with the materials of a design situation. Knowledge-Based Systems, 5(1), 3–14.

Weber, S. M. (2025). In Blickordnungen intervenieren: Organisationalen Wandel diskursanalytisch denken und gestalten. In M.-A. Heidelmann & S. M. Weber (Hrsg.), Organisationspädagogisches Denken und Forschen verstehen: Ein Studienbuch (S. 81–97). UTB; Waxmann.

Weber, S. M. & Heidelmann, M.-A. (2023). Images – Imagination – Imaginaries: Epistemic Organizing and Epistemologies of the Visual. In C. Cambre, E. Barromi-Perlman & D. Herman Jr. (Hrsg.), Visual Pedagogies (S. 37–57). BRILL. https://doi.org/10.1163/9789004529847_004

Weick, K. E. (1993). Organizational Redesign As Improvisation. In G. P. Huber & W. H. Glick (Hrsg.), Organizational Change And Redesign (S. 346–379). Oxford University Press.

Autorinnen und Autoren

Petra Gehring	petra.gehring@tu-darmstadt.de
Andreas Hebbel-Seeger	a.hebbel-seeger@macromedia.de
Dominikus Herzberg	dominikus.herzberg@mni.thm.de
Alinde Keller	keller@paedagogik.uni-kiel.de
Ines Langemeyer	ines.langemeyer@kit.edu
Christian Leineweber	christian.leineweber@ovgu.de
Antje Michel	antje.michel@fh-potsdam.de
Michael Pandey	michael.pandey@sz.uni-stuttgart.de
Gabi Reinmann	gabi.reinmann@uni-hamburg.de
Rüdiger Rhein	rhein@zqs.uni-hannover.de
Daniel Ruetten	d.ruetten@em.uni-frankfurt.de
Judith Simon	judith.simon@uni-hamburg.de
Isabel Steinhardt	isabel.steinhardt@uni-paderborn.de
Markus Stepanians	markus.stepanians@unibe.ch
Frank Vohle	frank.hubert.vohle@googlemail.com
Alice Watanabe	alice.watanabe@uni-hamburg.de
Nicolaus Wilder	wilder@paedagogik.uni-kiel.de
Silke Elisabeth Wrede	silke.wrede@fernuni-hagen.de
Dorothea Winter	dorothea.winter@uni-osnabrueck.de