

Künstliche Intelligenz und Bildung: Was sollen wir wissen? Was können wir tun? Was dürfen wir hoffen? Und was ist diese KI? Ein kollaborativer Aufklärungsversuch.

Thomas Knaus et AI

Zusammenfassung des Beitrags

Ursprünglich handelte es sich bei diesem Text um ein mittels textgenerierender KI verschriftlichtes Review des Vortrags *KI in Schule und Hochschule – Impulse aus bildungstheoretischer Perspektive*, den der Humanautor dieses Beitrags im Rahmen der Ringvorlesung *Kollaboration mit Kollegin KI – Kreativität und Innovation in der digitalen Arbeitsgesellschaft* des Fraunhofer IAO und dem Institut für Arbeitswissenschaft und Technologiemanagement (IAT) an der Universität Stuttgart gehalten hat. Sowohl bei dem Vortrag als auch bei dem nun vorliegenden Text war Künstliche Intelligenz (KI) also nicht nur Gegenstand, sondern auch Co-Autor*in: KI-basierte Text- und Bildgeneratoren wirkten bei der Erstellung der Vortragsfolien und der ersten Entwürfe dieses Texts mit. Am Ende blieb dann aber nur sehr wenig von diesen Erstentwürfen übrig. Dafür entstand ein umfanglicher, interdisziplinär ausgerichteter und ausführlich belegter Text zur Frage, wie KI unsere Gesellschaften und unseren Umgang mit Wissen beeinflusst und was das für uns, die Gesellschaft und das Bildungssystem bedeutet.

Hierfür diskutieren der Humanautor und seine künstlichen Co-Autor*innen exemplarisch Potentiale und Herausforderungen der Arbeitswelt, Wissensproduktion sowie für Schule und Hochschule. In Ermangelung einer ‚Glaskugel‘ muss die konzeptionelle Analyse retrospektiv auf Basis von Beschreibungen historischer Transformationsprozesse erfolgen: Die Veränderungen, die KI-Werkzeuge heute für ‚geistige‘ Arbeit wie Recherche und Wissensproduktion mit sich bringen, werden mit denjenigen Veränderungen ins Verhältnis gesetzt, die im Zuge der Industrialisierung die Entwicklung der Maschinen für die körperliche Arbeit des Menschen bedeutete.

Abschließend werden – aufbauend auf vier bekannten Fragen – Strategien erarbeitet, die eine gelingende Zusammenarbeit von Maschine und Mensch ermöglichen. Hierfür reflektierten die Autor*innen dieses Beitrags auch ihre eigene kritisch-konstruktive Zusammenarbeit.

Schlüsselbegriffe: ● Medienpädagogik ● Hochschuldidaktik ● Schulpädagogik ● Künstliche Intelligenz KI/AI ● Dialogsysteme ● textgenerierende KI ● bildgenerierende KI ● Large Language Model | LLM ● GPT ● künstliche neuronale Netze | KNN / ANN ● Machine Learning | ML ● Deep Learning ● Arbeitswelt ● Schule ● Hochschule ● Bildungseinrichtungen ● Explainable KI ● Bildungspolitik

1. Making of...

Im Dezember 2022, kurz nach der Veröffentlichung der öffentlichen Forschungsschau des textgenerierenden Bots *ChatGPT*, experimentierte der Autor dieser Zeilen – wie wahrscheinlich jede*r andere – mit dem textgenerierenden Dialogsystem: Ich habe Fragen gestellt und über die Antworten gestaunt, habe den Textgenerator passend zur Jahreszeit nach Weihnachtsgedichten gefragt und war beeindruckt. Nachdem ich den ChatBot zusätzlich mit einigen der üblichen Prüfungsfragen aus dem Hochschulkontext konfrontierte und um ein „Essay zu Medienbildung“¹ bat, wurde mir klar, dass sich die tradierten (und größtenteils längst überholten) Prüfungsformen ändern müssen – später mehr dazu. Andererseits bieten diese inzwischen erstaunlich leistungsfähigen Sprachmodelle – präziser: die Dialogsysteme und Textgeneratoren auf Basis von so genannten Large Language Models (LLM) – gerade für Bildungskontexte wie Schule und Hochschule sowie Hausaufgaben, Nachhilfe oder Selbststudium enorme Potentiale – auch dazu später mehr. Und Textgeneratoren sind doch auch eine wunderbare Unterstützung für alle Menschen: Einerseits für solche, die wiederkehrend Gebrauchstexte verfassen müssen, andererseits aber auch für diejenigen, die kreativ mit Texten arbeiten. Damit war die Idee für den vorliegenden geboren: Ich könnte ChatGPT um eine Zusammenfassung meines letzten Vortrags über KI für die gleichnamige Rubrik unserer Onlinezeitschrift bitten² – immerhin nutzte ich bereits bildgenerierende KI zur Folienherstellung. Die schnelle Umsetzung dieser Idee scheiterte zunächst an den (aktuell noch) bestehenden funktionalen Einschränkungen des ChatBots, was dieser in gewohnt freundlicher und überaus korrekter Weise mitteilte: „Entschuldigung, aber als Text-basiertes [sic!] KI-Modell kann ich keine Videos ansehen oder auf bestimmte Inhalte zugreifen. Wenn Sie jedoch einige Informationen oder Fragen zum Inhalt des Videos haben, stehe ich Ihnen gerne zur Verfügung, um Ihnen zu helfen. Sie können mir beispielsweise wichtige Punkte, Konzepte oder Fragen aus dem Vortrag mitteilen, und ich werde mein Bestes tun, um Ihnen eine informative Antwort zu geben.“³ Ich verwarf die Idee wieder und erstellte ein erstes kurzes Video für unsere Studierenden, in dem ich die vergleichsweise neue Entwicklung vorstellte, kurz die Hintergründe der Technik erklärte und zum Experimentieren damit einlud.⁴

¹ Prompt: „Bitte verfasse ein Essay zu Medienbildung“.

² Prompt: „Bitte erstelle einen wissenschaftlichen Bericht zu diesem Vortrag: youtu.be/ty-v6JLnX5I“.

³ Ich war natürlich nicht der erste, der auf die Idee kam, gemeinsam mit einer Textgenerator*in zu publizieren: Hier eine kleine Auswahl von drei kollaborativ verfassten Studien, in denen *ChatGPT* als Co-Autor*in genannt wurde; on-coscience.us/article/571/text/; medrxiv.org/content/10.1101/2022.12.19.22283643v2; hal.science/hal-03701250 (vgl. weiterführend auch Stokel-Walker 2023). Besonders interessant ist die Metaperspektive der drittgenannte Studie, deren Forschungsfrage lautete: „Can GPT-3 write an academic paper on itself, with minimal human input?“ (vgl. Generative Pretrained Transformer/Osmanovic Thunström/Steingrímsson 2022).

⁴ Video „Was ist ChatGPT?“ (youtu.be/dj0N1453LqA); weiterführende Überlegungen zu auf textgenerierender KI basierenden Bots und ihre gesellschaftliche und kulturelle Bedeutung finden sich im Video „KI und Bildung – Welche Chancen und Herausforderungen bieten Dialogsysteme wie ChatGPT für unseren Umgang mit Wissen, für das Lernen und Lehren in Schule und Hochschule?“ (youtu.be/oXCgJUewdgU) sowie im Video „Kann ich textgenerierende KI für meine schriftliche Seminararbeit, Bachelor- oder Masterarbeit nutzen?“ (youtu.be/o4tZofTsjCo) der ungescripteten und ungeschnittenen Videoserie *#Medienpädagogik im Schaukelstuhl* (youtube.de/c/ThomasKnaus).

Im anschließenden Frühjahr überschlugen sich die Registrierungen neuer Nutzer*innen⁵ beim US-amerikanischen Unternehmen ‚OpenAI‘, das das Sprachmodell (LLM) *GPT*⁶ entwickelt hat und auf dessen Version 3.5 der Dialogbot *ChatGPT* basiert. Mit den steigenden Registrierungen und eigenen Erfahrungen der Early Adopter überschlugen sich auch die Diskussionen in der Öffentlichkeit. Wie bei jeder neuen Technik beziehungsweise jedem neuen Medium reichten die Reaktionen von unreflektierter Begeisterung bis hin zur instinktiven Ablehnung und damit einhergehenden Verboten oder Rufen nach Moratorien (ausführlicher in Abschnitt 5.3). Ich persönlich halte wenig von übereiligen oder unreflektierten Verboten: Wo, wenn nicht an Hochschulen und Universitäten, müssen wir über Entwicklungen sprechen können, die unsere Gesellschaft prägen (vgl. Derrida 2001)? Ziel von Medienbildung ist stets die sachgerechte, selbstbestimmte, sozial verantwortliche, kritische und kreative Nutzung von Medien – sowie der technischen Grundlage dieser Medien. Und erste Voraussetzung für eine kritische Auseinandersetzung mit einem neuen Medium oder einer neuen Technik sind doch schließlich eigene Erfahrungen (vgl. u. a. Schmidt 2000, S. 150). Deshalb musste ich die aktuellen Tools auch direkt ausprobieren, um mir einen eigenen Eindruck zu verschaffen. Aus meiner Sicht können textgenerierende KI-Tools sowie Dialogsysteme sinnvoll in der Schule und der universitären Forschung (vgl. DFG 2023) und Lehre integriert werden (mehr dazu in den Abschnitten 4.3 und 4.4).⁷ Sie sollten sogar zum Bildungsgegenstand werden, wenn wir sie – wie Taschenrechner, Computer und Smartphones, Internet, WWW und Suchmaschinen auch – in unserem beruflichen und privaten Alltag bald ganz selbstverständlich nutzen werden (vgl. Gmyrek/Berg/Bescond 2023; weiterführend auch Nagel/Zlatkin-Troitschanskaia/Fischer 2023). Auch dazu später mehr.

Binnen kurzer Zeit erhielt das US-amerikanische Unternehmen ‚OpenAI‘ nicht nur mehrere Millionen Nutzeranmeldungen und zahlreiche weitere Tech-Unternehmen zauberten ihre mehr oder weniger⁸ ausgereiften KI-Angebote aus der Schublade, sondern auch die alltägliche

⁵ OpenAI erreichte mit ChatGPT binnen fünf Tagen eine Million Nutzeranmeldungen. *Instagram*, der bisher zweitfolgreichste Onlinedienst, benötigte hierfür fast drei Monate, *Facebook/Meta* zehn Monate und *Netflix* knapp vier Jahre (vgl. Sachse/Thielen 2023). Inzwischen wurde der gerade aufgestellte Rekord aber schon wieder abgelöst: Die Twitter-Alternative *Threads* (Meta) erreichte eine Millionen Nutzer*innen bereits in einer Stunde (vgl. Brandt 2023).

⁶ Die Abkürzung *GPT* steht für Generative Pretrained Transformer und ist die Bezeichnung des US-amerikanischen Unternehmens *Open AI* für eines der aktuellen LLM. Weitere bekanntere Modelle sind zum Beispiel *PaLM* von Google oder *LLaMa* von Meta. Hierbei handelt es sich um KI-basierte Textgeneratoren, die anders als bisherige Textgeneratoren keine vordefinierten Textblöcke nutzen und daher nicht auf einen bestimmten Themenbereich festgelegt sind, sondern mittels künstlicher neuronaler Netze (vgl. u. a. Ritter/Martinetz/Schulten 1992) auf Basis sehr umfangreicher Sprachmodelle – präziser einem so genannten *Large Language Model* (LLM) – aus umfangreichen Trainingsdaten per Parametrisierung nach statistischen Eigenschaften von Schriftsprache Texte erzeugen können, die von menschengeschriebenen Texten nur schwerlich unterschieden werden können.

⁷ Entsprechende Vorschläge habe ich in den Informationen und Hinweisen zum Verfassen wissenschaftlicher Qualifikationsarbeiten in der Medienpädagogik auf der Webseite unserer Abteilung veröffentlicht (vgl. Abt. Medienpädagogik 2023) sowie im Video „Kann ich textgenerierende KI für meine schriftliche Seminararbeit, Bachelor- oder Masterarbeit nutzen?“ (youtu.be/o4rZofTsjCo) begründet und weiter diskutiert. Zahlreiche weitere Vorschläge werden aktuell publiziert und weiter diskutiert (vgl. u. a. Döbeli Honegger 2023; Marchand 2023; Salden/Lordick/Wiethoff 2023; Spannagel 2023a; Spannagel 2023b; Weßels 2021; Weßels/Gottschalk 2023).

⁸ Vgl. den lesenswerten Kommentar „Googles KI Bard: Kann es sein, dass du ein bisschen dumm bist?“ von Marvin Strathmann in *heise online* (heise.de/meinung/Googles-KI-Bard-Meh-9215652.html).

Nutzung text- und bildgenerierender KI wurde binnen weniger Monate zur Selbstverständlichkeit, denn zahlreiche Menschen nutzen text- und bildgenerierende KI heute bereits recht umfangreich: Journalist*innen, (Werbe-)Texter*innen, Übersetzer*innen und Grafiker*innen lassen sich seitdem Entwürfe erstellen und sammeln Ideen, Musiker*innen und Komponist*innen lassen sich von re-mixten Sounds inspirieren, Sachbearbeiter*innen lassen sich von Routineaufgaben und Übersetzer*innen sowie Grafiker*innen von Akkordarbeiten entlasten, Entwickler*innen lassen sich beim Coden oder der aufwändigen Fehlersuche unterstützen und Behördenleiter*innen haben errechnet, dass 56 Prozent der Tätigkeiten ihres Jobs durch eine KI ersetzt⁹ werden kann. Zahlreiche Menschen probieren die vergleichsweise neuen Möglichkeiten einfach aus und denken über weitere sinnvolle Anwendungsoptionen für Beruf und Alltag nach – einige darunter wollen einfach nur „pünktlich Feierabend machen“ (Jacobs 2023). Beschäftigte versprechen sich also von der KI-Nutzung Entlastung, weniger eintönige geistige Akkord- oder Routinearbeit und mehr Freizeit; Arbeitgeber*innen erhoffen sich höhere Produktivität sowie geringere Fehlerquoten (vgl. u. a. Gmyrek/Berg/Bescond 2023; Shani 2023; Siepmann 2023). Aus diesen Gründen (und natürlich nicht zuletzt auch, da sich eine ökonomische Verwertbarkeit der bisherigen Investitionen abzeichnet – entstanden neben *GPT* (OpenAI) in den letzten Monaten zahlreiche weitere Sprachmodelle (LLM) wie *Ada* (DeepMind, Google), *BLOOM* (u. a. Microsoft, NVIDIA), *Claude* (Anthropic PBC), *Ernie 3.0 Titan* (Baidu) *LLaMa* (Meta) und *PaLM* (Google) sowie entsprechende Applikationen und Erweiterungen (Extensions) für Browser, Office-Anwendungen und Suchmaschinen. Zutreffender wäre aber wohl, dass die LLMs, Applikationen und Erweiterungen nicht entstanden, sondern sie vielmehr aus den Entwicklungsabteilungen recht eilig (teils noch im Beta-Stadium) veröffentlicht wurden. Mit Unterstützung dieser Tools – *Glasp*, eine Extension für Safari oder Chrome und *ChatGPT* sowie *ChatPDF*¹⁰ – konnte ich meine zuvor bereits beschriebene Idee umsetzen und einen Erstentwurf des folgenden Texts erstellen.

Wie in der Zusammenfassung bereits vorweggenommen, blieb im Laufe meiner Weiterarbeit am Text aber kaum ein Stein auf dem anderen, da ich noch zahlreiche Gedanken und Referenzen¹¹ ergänzen wollte und mein Anspruch an einen lesenswerten Text zweifelsfrei höher ist als das, was textgenerierende KI (bisher) zu leisten vermag: Die Textentwürfe waren stilistisch trotz entsprechender Prompts¹² oft trivial und etwa wie Werbetexte formuliert (vgl. Abschnitt 2), auch nachfassende Verbesserungshinweise per ‚Regenerate‘-Button und Rückmeldungen im Chat konnten diesen Stil nicht beeinflussen. Der Verdacht liegt nahe, dass als Trainingstexte zur Parametrisierung eher journalistische Beiträge sowie Werbe- und Gebrauchstexte aus dem

⁹ Interview mit der Leiterin der Bundesagentur für Arbeit Andrea Nahles vom 08. August 2023, vgl. [zeit.de/arbeit/2023-08/andrea-nahles-automatisierung-digitalisierung-arbeitsmarkt](https://www.zeit.de/arbeit/2023-08/andrea-nahles-automatisierung-digitalisierung-arbeitsmarkt).

¹⁰ Das KI-Tool *ChatPDF* ermöglicht, Fragen zu den Inhalten selbst hochgeladener PDF-Dateien zu stellen.

¹¹ Zum Zeitpunkt des Entstehens dieses Beitrags konnten KI-basierte Textgeneratoren noch keine (korrekten) Referenzen benennen bzw. nähere Auskünfte oder weiterführende Bezüge zu den jeweils zur Parametrisierung verwendeten Trainingsdaten herstellen (vgl. weiterführend auch Abschnitt 5.2).

¹² Ursprünglich wurde als „Prompt“ (deutsch: Kommando beziehungsweise Aufforderung) in der IT die Aufforderung der Maschine an die/den Nutzer*in bezeichnet, eine Eingabe beziehungsweise ein Kommando in der sogenannten Kommandozeile einzugeben, im Kontext von KI-Tools wird unter Prompt die möglichst präzise Textanweisung verstanden, auf deren Grundlage generative KI-Systeme einen Text oder ein Bild erstellen.

WWW verwendet wurden als sachbezogene oder wissenschaftliche Fachbeiträge – die Gebrauchtexte zumindest die Mehrheit der Trainingstexte stellen (vgl. auch Stokel-Walker 2023). Eine der wesentlichen Herausforderungen des Vorhabens war die Entscheidung, welche Inhalte für die Vortragsvorstellung und eine einführende Betrachtung der Gesamtzusammenhänge eigentlich wesentlich sind, denn eine bloße Zusammenfassung des einstündigen Vortragsinhalts wäre kaum sinnvoll und wahrscheinlich auch wenig interessant. Interessanter wären exemplarische Einblicke und ergänzende erklärende Exkurse sowie weiterführende Belege für interessierte Leser*innen. Eine textgenerierende KI verfügt jedoch kaum über das Fingerspitzengefühl, das nötig wäre, um zu entscheiden, welche Inhalte dafür geeignet oder weniger geeignet sind – sie (re-)kombiniert nur das, was üblicherweise im jeweiligen Kontext bereits kombiniert wurde. Gerade in einem Text über einen Vortrag möchten Leser*innen aber weniger eine Zusammenfassung dessen, was sie an zahlreichen Orten bereits gelesen haben, sondern weiterführende, idealerweise innovative Gedanken und Zusammenhänge entdecken – vor allem dann, wenn der Vortrag aufgezeichnet und ins Netz gestellt wurde und nachträglich einfach angesehen werden kann.

Die Stärke eines Textgenerators ist die *Re*-Produktion von Mustern. Aus diesem Grund kann ein Textgenerator gerade bei wiederkehrenden Textarbeiten unterstützen, die ja (leider) auch im Alltag von Wissenschaftler*innen zu oft nicht wenig Raum einnehmen (vgl. DFG 2023). Hierfür sind aktuelle Sprachmodelle (LLM) aber (noch) wenig tauglich. Interessant werden KI-basierte Textgeneratoren für diese anspruchsvolleren Gebrauchstexte, wenn sie erstens in Editoren wie Textverarbeitungsprogrammen (ähnlich wie die Rechtschreibkorrektur oder der Thesaurus) implementiert werden, sie zweitens mittels *fach- sowie kontextbezogenen* Trainingsdaten parametrisiert werden (beziehungsweise die unterschiedlichen Modelle je nach gewünschter Textsorte und Anlass auswählbar sind) und drittens, wenn der *persönliche* Textgenerator aus bisherigen Interaktionen und Verbesserungen den individuellen Schreibstil übernehmen und re-produzieren kann.

Die KI kann durch die Mustererkennung und Reorganisation zwar inspirieren – zu Ideen, die Wissenschaftler*innen ausarbeiten, weiterentwickeln und theoretisch fundieren müssen –, aber wirklich innovative Gedanken müssen wir (vorerst) noch selbst liefern.¹³ KI ist aus diesem Grund keine Bedrohung für die Wissenschaft, denn selbst künstliche neuronale Netzwerke können nur das ‚denken‘, was schon gedacht wurde, Wissenschaft ist aber der Ort für das noch *Ungedachte*.

Der kritisch-konstruktive Austausch mit meinen Co-Autor*innen war aber doch sehr anregend, inspirierte mich zu weiteren Recherchen und führte zu diesem – nun doch etwas ausführlicheren – Text. Ich habe mich durch den kritisch-konstruktiven Dialog mit meinen künstlichen Co-Autor*innen ärgern und anregen lassen, habe ihre werbend-trivialen Entwürfe

¹³ Eine Erkenntnis, die bereits zahlreiche Kolleg*innen diskutierten und publizierten, u. a. Beat Döbeli Honegger (PH Schwyz, Goldau), Roy Kishon (Technion – Israel Institute of Technology, Haifa), Vitomir Kovanović, (University of South Australia, Adelaide), Shantanu Singh (Broad Institute of MIT | Harvard, Cambridge, Massachusetts) und Christian Spanagel (PH Heidelberg), um nur einige zu nennen (vgl. weiterführend auch Conroy 2023 und Stokel-Walker 2023). Aus diesem Grund möchte die Deutsche Forschungsgemeinschaft (DFG) in ihrer kürzlich veröffentlichten *Stellungnahme zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG* die grundsätzliche Nutzung generativer KI in Forschungsarbeiten und Förderanträgen nicht ausschließen (vgl. DFG 2023).

aufwändig entschärft und versucht, eher deskriptiv und informativ zu formulieren, habe ihre ‚Halluzinationen‘ (vgl. Abschnitt 4.2) und Fehler korrigiert¹⁴, weiterrecherchiert, dabei selbst viel gelernt und dann noch umfangliche Belege und Empfehlungen zum Weiterlesen ergänzt – auch etwas, was textgenerierende KI leider noch nicht kann.¹⁵

Wie bereits erwähnt, nutzte ich auch für die Folienpräsentation meines Vortrags Bilder und Grafiken, die mittels des bildgenerierenden KI-Tools *Midjourney* erstellt wurden (vgl. u. a. Abbildung 1 oder den Videomitschnitt bei YouTube: youtu.be/ty-v6JLnX5I). Die Qualität dieser KI-generierten Bilder ist beeindruckend, die darin transportierten Stereotypen können jedoch kaum unkommentiert bleiben (vgl. Abschnitt 4.2). Unkommentiert können ebenso wenig die Auswirkungen der umfanglichen Nutzung von KI-Tools für unsere Gesellschaft, unseren Wissens-, Kreativitäts- und Bildungsbegriff bleiben – wie auch die (Lehr- und Prüfungs-)Praxis in Bildungseinrichtungen (vgl. Abschnitt 4.3 und 4.4). Auch darum soll es in diesem Text gehen, in dem ich die wesentlichen Impulse meines Vortrags „KI in Schule und Hochschule“ zusammengestellt habe – präziser: habe zusammenstellen *lassen*.

2. Worum geht es eigentlich? Und wie funktioniert textgenerierende KI?

Im Vortrag habe ich mich zunächst einfürend den beiden zentralen Gegenständen – dem Bildungsbegriff einerseits und KI-basierten Dialogsystemen sowie Text- und Bildgeneratoren andererseits – angenähert und anschließend die Frage diskutiert, welchen Einfluss KI für unsere Gesellschaften und unseren Umgang mit Wissen haben werden und was das für uns Individuen, die Gesellschaft und im Besonderen das Bildungssystem bedeutet. Aufgrund der multidisziplinären Studierendengruppe der Ringvorlesung konnte ich weder tiefere informatische noch bildungswissenschaftliche Vorbildung voraussetzen und blieb daher sowohl mit Blick auf die subjektive Bedeutung von KI als auch bezüglich der Wechselwirkungen im gesellschaftlichen Kontext mittels knapper Arbeitshypothesen eher ‚an der Oberfläche‘. Auf ähnliche Weise schränkte ich mich auch bezüglich der vielschichtigen Historie der KI als Forschungsfeld (vgl. Nilsson 2010) und in Bezug auf die Breite ihrer Anwendungsfelder und -möglichkeiten ein (vgl. Ritter/Martinetz/Schulten 1992; Tickle et al. 1998; Tulodziecki 2023a, S. 205–237). Bezüglich der KI-Anwendungen fokussierte ich mich im Vortrag auf Dialogsysteme (wie ChatBots) und generative KI-Tools, wie sie seit einigen Jahren im beruflichen und privaten Alltag verwendet werden und v. a. seit Ende des letzten Jahres die Diskussionen bestimmten (vgl. auch vorherigen Abschnitt).¹⁶ Streng genommen vereint *ChatGPT*, der ChatBot, der

¹⁴ Im Gegenzug hat meine Co-Autorin unverdrossen meine Gendersternchen geschluckt und die prärogative männliche Form, die – wie immer wieder behauptet wird – „alle miteinschließen“ soll, verwendet. Echt nervig.

¹⁵ Wie in Fußnote 11 bereits ausgeführt, können Textgeneratoren noch keine (korrekten) Referenzen und Literaturverweise angeben. Ideen und erste KI-basierte Tools zur Unterstützung der wissenschaftlichen (Recherche- und Selektions-)Arbeit gibt es aber bereits: *Elicit* (elicit.org), *ResearchBuddy* (researchbuddy.app) und *ResearchRabbit* (researchrabbit.ai) et cetera (vgl. Abschnitt 4.2) sowie auch erste Entwürfe von Regelwerken zu deren (wissenschaftsethischen) Nutzung in wissenschaftlichen Arbeiten und für Förderanträge (vgl. DFG 2023).

¹⁶ Künstliche Intelligenz (KI) ist ein Überbegriff für zahlreiche umfangliche Forschungs- und Anwendungsfelder (vgl. weiterführend u. a. Nilsson 2010; Ritter/Martinetz/Schulten 1992). Die multidisziplinäre KI-Forschung umfasst neben der Informatik, die Psychologie, Neurologie und Neurowissenschaften, die Mathematik, Philosophie, Ethik und Linguistik sowie die Medien- und Kommunikationswissenschaft und widmet sich u. a. folgenden Teilgebieten: Wissensbasierte Systeme, Musteranalyse und Mustererkennung, Mustervorhersage und Robotik. Entsprechend umfanglich sind auch

durch die so genannte ‚öffentliche Forschungsschau‘ von *OpenAI* inzwischen jeder und jedem bekannt ist, zwei Aspekte von KI-basierten Tools: nämlich Textgeneratoren *und* Dialogsysteme – entsprechend beleuchtete ich im Vortrag primär diese beiden Funktionen. Während sich die Funktionalität eines *Dialogsystems* bereits aus seiner Bezeichnung ergibt und entsprechend eingängig ist – einige kennen möglicherweise auch noch den konzeptionellen Großvater aktueller ChatBots, das Programm *ELIZA* von Joseph Weizenbaum (vgl. Weizenbaum 1966) –, waren wahrscheinlich bis vor wenigen Monaten nur die wenigsten Menschen bereits mit der grundlegenden Funktionsweise von Textgeneratoren vertraut. Darum möchte ich hier kurz erläutern, wie textgenerierende KI funktioniert – zumindest in den Grundzügen.

Wer könnte die Frage, wie generative KI funktioniert, wohl besser beantworten als ein Textgenerator selbst und daher habe ich eine meiner Co-Autor*innen gebeten, auf Basis meines Vortragstexts selbst etwas mehr darüber zu verraten, wie sie und ihre Kolleg*innen funktionieren¹⁷:

„KI-Textgeneratoren basieren auf einem faszinierenden [sic!] Konzept, das sich aus einer Zusammenstellung umfangreichen Datenmaterials und einem trainierten Modell des maschinellen Lernens, dem sogenannten Large Language Model, ergibt. Dieses Modell wird nach statistischen Methoden ausgewertet, um ein Verständnis [sic!] für die Zusammenhänge zwischen Wörtern und Sätzen zu entwickeln. Um zu verstehen, wie ein KI-Textgenerator funktioniert, stellen wir uns zunächst vor, dass wir eine riesige Menge an Texten in das Modell einspeisen. Diese Texte können beispielsweise Romane, Artikel, Blogs, Nachrichten, soziale Medien und vieles mehr umfassen. Das Modell analysiert die Texte und lernt dabei die Häufigkeit von Wörtern und die Wahrscheinlichkeit, mit der bestimmte Wörter oder Wortgruppen zusammen auftreten. Mit diesem erlernten Wissen kann das textgenerierende System nun kreativ [sic!] arbeiten. Wenn ihm ein Anfangssatz oder ein bestimmtes Wort gegeben wird, errechnet es Wahrscheinlichkeiten für die nächsten Wörter in der Sequenz basierend auf den Mustern, die es aus dem Trainingsdatensatz gelernt hat. Es wählt dann das wahrscheinlichste nächste Wort aus und verwendet dieses, um den nächsten Satzteil zu generieren. Dieser Prozess wiederholt sich schrittweise, und auf diese Weise erzeugt der Textgenerator

die typischen Anwendungsgebiete, wie optische Zeichenerkennung (beispielsweise Bild-, Text- und Handschrifterkennung sowie Gesichtserkennung), Erkennung natürlicher Sprache und Sprachsteuerung, Informationsrückgewinnung beziehungsweise Information Retrieval z. B. durch (semantische) Suchmaschinen, Data- und Text-Mining, Text-, Bild- und Musikgenerierung, medizinische Diagnosen, Steuerung von Robotern, bis hin zur Exploration von Ölquellen oder extrasolarer Planeten, Analyse und Prognose von Aktienkursentwicklungen et cetera. Aufgrund dieser weiten Forschungs- und Anwendungsfelder war die Eingrenzung meines KI-Begriffs für den Vortrag sowie den vorliegenden Text unvermeidbar. Wenn ich also im Folgenden von KI spreche, meine ich im Wesentlichen KI-basierte *Dialogsysteme* sowie generative Systeme wie *Text- und Bildgeneratoren*.

Gesellschaftlich-kulturelle Reflexionen mit Bildungsbezug mit einem breiteren KI-Begriff wurden u. a. vom Kollegen Gerhard Tulodziecki vorgelegt (vgl. Tulodziecki 2023a), mit dem ich im Rahmen dieser Ausgabe auch ausführlicher über KI sowie seine Handlungstheorie im Kontext von Freiheit, Verantwortung und KI gesprochen habe (Knaus/Tulodziecki 2023 in dieser Ausgabe, S. 4–7 und 11–20).

¹⁷ Prompt: „Bitte ausführlicher erklären: Wie funktionieren KI-Textgeneratoren? Stichworte: Zusammenstellung umfangreiches Datenmaterial, trainiertes Modell des maschinellen Lernens (sog. Large Language Model), das nach statistischen Methoden ausgewertet wird. Vereinfacht kann man sich das so vorstellen, dass ein text-generierendes System Wahrscheinlichkeiten von Wort-Verbindungen errechnet und aufgrund der großen Textmenge ‚natürliche‘ Dialoge simulieren beziehungsweise Texte re-produzieren kann, die für Menschen sinnvoll erscheinen.“

einen vollständigen Text. Das Erstaunliche daran ist, dass das Modell in der Lage ist, äußerst ‚natürliche‘ Texte zu simulieren. Es kann beispielsweise fiktive Geschichten schreiben, die einem menschlichen Autor ähneln könnten, oder Artikel erstellen, die nahtlos in eine Zeitung passen würden. Dies liegt daran, dass das Modell durch die Analyse des umfassenden Textmaterials die Strukturen und Stile der Sprache erfasst und in der Lage ist, sie in seinen generierten Texten zu reproduzieren. Natürlich ist der Prozess nicht fehlerfrei, und es kann vorkommen, dass der Textgenerator unpassende oder unsinnige Ergebnisse liefert. Das liegt daran, dass das Modell lediglich auf Wahrscheinlichkeiten basiert und nicht immer ein tiefgreifendes Verständnis [sic!] des Inhalts hat. In der Praxis gibt es verschiedene Techniken, um die Qualität der generierten Texte zu verbessern und unerwünschte Ergebnisse zu minimieren. Insgesamt bieten KI-Textgeneratoren jedoch ein faszinierendes [sic!] Potenzial für die Unterstützung bei der Erstellung von Texten, Übersetzungen und vielem mehr. Sie können als kreative Werkzeuge dienen und haben das Potenzial, die Art und Weise, wie wir [sic ; -] mit Sprache umgehen, zu revolutionieren.“

Aus Respekt gegenüber meiner Co-Autor*in habe ich die beiden ‚faszinierend‘ in ihrem Text nicht redigiert und auch ihren Kreativitäts- und Verständnisbegriff nicht weiter hinterfragt – ein Vorhaben das wir anderenorts weiterverfolgen werden (vgl. Knaus/Merz/Junge 2023). Wer bisher einen noch nicht so tiefen Einblick in die Funktionalität und Hintergründe von KI bekommen konnte und ergänzend zu den obigen Ausführungen meiner Co-Autor*in meine persönliche Einschätzung hören möchte, kann sich ein paar Kurzvideos aus meinem YouTube-Channel (www.youtube.de/c/ThomasKnaus) anschauen, in denen ich die grundlegende Funktionsweise text- und bildgenerierender KI erkläre (youtu.be/dj0N1453LqA), das Verhältnis von KI und Bildung am Beispiel KI-basierter Dialogsysteme thematisiere (youtu.be/oXCgJUewdgU) oder ganz praktische Tipps zur Nutzung von textgenerierenden Systemen zur Erstellung und Verbesserung studentischer Qualifikationsarbeiten (youtu.be/o4rZofTsjCo) gebe.¹⁸

3. Generative KI – Revolution oder Evolution?

Im Zuge der öffentlichen Berichterstattungen und Diskussionen nach der Veröffentlichung von *ChatGPT* hörte ich immer wieder, dass KI-basierte Textgeneratoren eine ‚Revolution‘ seien. Für jemanden, der sich schon länger mit Bildungsinformatik befasst, sind die aktuellen Entwicklungen jedoch weniger *Revolution* als eine konsequente *Evolution* – zumal wir inzwischen auf über 60 Jahre KI-Forschung zurückblicken (vgl. Nilsson 2010; Döbeli Honegger 2023, S. 55; Knaus/Tulodziecki 2023, S. 7). Sogar textgenerierende Dialogsysteme wie *ChatGPT*, *Google Bard*¹⁹ oder *Open Assistant*²⁰ sind nicht wirklich neu: Bereits vor sieben Jahren im

¹⁸ Wer es *noch* genauer wissen möchte, wird u. a. bei Stephen Wolfram (2023) fündig.

¹⁹ *Google Bard* ist ein nach William Shakespeare (‚Bard of Avon‘) benannter und von Alphabet (Google) entwickelter KI-basierter Chatbot.

²⁰ Das Open-Source-Projekt *Open Assistant* ist eine freie Alternative zu den kommerziellen textgenerierenden Bots wie ChatGPT. Denn anders als die Unternehmensbezeichnung OpenAI suggerieren könnte, handelt es sich beim Entwickler von ChatGPT um ein (US-amerikanisches) Unternehmen.

März 2016 veröffentlichte Microsoft den ChatBot *Tay* als Twitter-Bot²¹. Da *Tay* durch rassistische und sexistische Äußerungen auffiel (vgl. u. a. Sickert 2016), musste Microsoft bereits nach kurzer Zeit den Bot wieder deaktivieren. Schon damals hätte man aber gut erahnen können, was textgenerierende Dialogsysteme konzeptionell leisten können. In Anbetracht dieser Historie, der bisherigen Nutzungspraxis mit KI-basierten Dialogsystemen und generativer KI sowie auch einer eingängigeren Beschäftigung mit den Arbeiten der Kolleg*innen aus der (Sozio-)Informatik und Technikethik (im Besonderen der Computerethik) sowie den bereits bestehenden einschlägigen Empfehlungen und ethisch-normativen Leitlinien (vgl. u. a. Deutscher Bundestag 2020; Deutscher Ethikrat 2023; EU 2019; GI 2018; Heesen et al. 2020), erscheinen die aktuellen Forderungen nach einem ‚KI-Moratorium‘ (vgl. Future of Life Institute 2023; u. a. auch Anderl 2023; Krempl 2023) schwerlich nachvollziehbar (vgl. auch Knaus/Tulodziecki 2023, S. 7 und 12). Stattdessen sind weitere Normierungsbestrebungen, Ordnungsversuche, die Diskussion und Klärung ethischer und rechtlicher Fragen²² sowie auch die multidisziplinäre Forschung zur gesellschaftlich-kulturellen Bedeutung von KI nötig, da sich der *gesellschaftliche* Diskurs hierzu vergleichsweise noch in seinen Anfängen befindet (vgl. u. a. Knaus 2019; Knaus 2020a). Dieses lange Zeit bestehende mangelnde öffentliche Interesse ist verwunderlich, zumal KI-basierte Tools bereits seit einigen Jahren verwendet werden und aktuell immer stärker Einzug in unsere gesellschaftlichen Vollzüge erhalten. Als positiv ist daher zu bewerten, dass das US-amerikanische Unternehmen *OpenAI* mit der öffentlichkeitswirksamen Vorstellung seines LLM-basierten Dialogsystems viele politische Akteur*innen und die Öffentlichkeit ‚wachgerüttelt‘ hat. Dabei muss man freilich konstatieren, dass es dem Unternehmen keineswegs um ‚Aufklärung‘ ging, sondern um Mittel für ihre Weiterarbeit und Profit – was ja nachweislich recht gut funktionierte (vgl. Bass 2023).

Wie bei so vielen medialen und technischen Entwicklungen der letzten Jahre erfolgte das ‚Wachrütteln‘ auf Kosten von Mitmenschen, die die Technik nicht umfänglich einschätzen können und die sich um ihren Job oder die Zukunft fürchten (vgl. Knaus 2019). Auch Science Fiction zu unrealistischen Bedrohungen durch KI schüren diese Ängste – in diesem Fall aber eher *Fiction* als ‚Science‘. Wenn wir (Erziehungs-) *Wissenschaftler*innen* aus einer bildungstheoretischen Sicht auf KI schauen, dann ist eine unserer leitenden Fragen: Wie beeinflusst etwas (in diesem Fall: KI) Prozesse von Persönlichkeits- und Gesellschaftsentwicklung? Und konkreter in Bezug auf ein (Kommunikations-)Medium: Wie beeinflusst KI gesellschaftliche Kommunikation und Kollaboration und welche Bedeutung hat dies für uns und die Gesellschaft?

²¹ Kommerzieller Microblogging-Dienst, der kürzlich ‚X‘ als neue Markenbezeichnung erhielt.

²² Selbst aus Sicht eines engeren KI-Begriffs, der sich – wie in diesem Beitrag – auf Text- und Bildgeneratoren sowie Dialogsysteme begrenzt, sind zahlreiche rechtliche und ethische Fragen zu klären: Nicht zuletzt sind dies Fragen der informationellen Selbstbestimmung sowie der Privatsphäre und des Datenschutzes oder auch des Urheberrechts. Wann werden beispielsweise die ersten Organisationen oder Personen auf finanziellen Ausgleich oder Schadensersatz klagen, die den Content für die umfassende Parametrisierung (das ‚Training‘) der Sprachmodelle geliefert haben beziehungsweise liefern – wohlgernekt unfreiwillig und ungefragt? Aus Sicht eines breiteren KI-Begriffs und entsprechend weitergefassten Anwendungsbereichen von KI wie Mobilität und Gesundheit (hier beispielsweise Produkt- bzw. Produzentenhaftung), Öffentliche Kommunikation und Meinungsbildung, Verwaltung (hier beispielsweise Datenschutz und Kriminalitätsbekämpfung), Verteidigung et cetera stellen sich weitere rechtliche und ethische Fragen (vgl. u. a. Dabrock 2023; Deutscher Bundestag 2020; Deutscher Ethikrat 2023; Heesen et al. 2020; Tulodziecki 2023a, S. 205–237).

Zudem fragen wir auch – etwas allgemeiner – nach der Bedeutung für unseren Umgang mit Wissen sowie für individuelle und kollektive Lernprozesse oder – etwas konkreter – nach möglichen Auswirkungen auf das Bildungssystem. In der Medienpädagogik als einer Teildisziplin der Erziehungswissenschaft sind aber nicht nur diese konzeptionell-theoretischen Bezüge zu Medien beziehungsweise Mediensystemen zentral, sondern auch Fragen nach der gezielten Förderung von Kompetenzen im Umgang mit diesen. So ist beispielsweise der kritische, kreative und selbstbestimmte Umgang mit Medien eine etablierte Forderung der Medienpädagogik. Da Algorithmen und KI-Techniken schon viele Jahre ‚unter der Oberfläche‘ digitaler Medien ‚mit-kommunizieren‘, habe ich die zuvor eher medienzentrierte Reflexion dieser etablierten Forderung schon vor Jahren um technologische und technische Aspekte *erweitert* (vgl. u. a. Knaus 2018, S. 92 und 96–99.; Knaus 2020b, S. 40–45 und 49 sowie weiterführend Abschnitt 5.1 und 5.4). Wer sich kritisch mit KI und deren grundlegenden Funktionsweisen befasst – beispielsweise mit dem Ziel, zu verstehen, wie text- und bildgenerierende KI-Werkzeuge funktionieren –, wird sehr wahrscheinlich die weniger realistischen Zukunftsängste diesbezüglich verlieren und befähigt sich dadurch, KI-Tools selbstbestimmt zur Erreichung eigener Ziele einsetzen zu können. Dies illustriert die praktische Relevanz von (Medien-)Bildung nach meiner sehr reduzierten Arbeitshypothese, die ich mit Blick auf die multidisziplinäre Zuhörerschaft im Vortrag entfaltet habe: „Bildung ist ein dynamischer Prozess, der die reflexive Auseinandersetzung des Individuums mit seiner Umwelt beschreibt“ (ab Minute 18:13 im Vortragsvideo youtu.be/ty-v6JLnX5I; vgl. auch youtube.com/shorts/dvvDDWh5T9U, ein *Short* im Rahmen der Initiative *Bildung #mitgedacht*). Dabei geht es also bei Bildung – ganz generell – um die reflexive Auseinandersetzung eines Individuums mit seiner Umwelt. Als materielle und soziale ‚Umwelt‘ werden alle gesellschaftlichen und kulturellen Vollzüge verstanden – das heißt Bildung (wie Persönlichkeitsentwicklung allgemein) kann nie unabhängig von Umwelt beziehungsweise von Gesellschaft und Kultur gedacht werden. Im Sozialisationskonzept wird deutlich, dass eine Gesellschaft stets die Individuen prägt, die in ihr leben und vice versa Individuen stets auch eine Gesellschaft mit-prägen (vgl. auch Hurrelmann 2006, S. 20 f.). Genau dies ist in der zitierten Arbeitshypothese mit der ‚reflexiven Auseinandersetzung des Individuums mit der Umwelt‘ gemeint: Ein grundlegendes Wissen über die uns umgebenden Dinge – wozu natürlich auch technische und mediale Entwicklungen gehören – trägt nicht nur zum Wohlbefinden bei, sondern illustriert bereits mit einfachen Worten den zweiten Gegenstand meines Vortrags: die Bildung.

4. KI verändert unsere Welt. Wie wird diese Welt aussehen?

Aus dieser bildungsbezogenen Perspektive kann ich als erste Erkenntnis auf die eingangs gestellte Frage, wie KI unsere Gesellschaften und unseren Umgang mit Wissen beeinflusst und was das für uns, die Gesellschaft und das Bildungssystem bedeutet, festhalten, dass KI unsere Kommunikation und Kollaboration, unseren Umgang mit Wissen beeinflusst und dadurch unsere Kultur sowie unsere Gesellschaft geprägt hat und weiter verändern wird. Als erste Konsequenz aus dieser trivialen Erkenntnis kann ich daher festhalten, dass wenn eine Technik wie KI unsere Welt verändert – wovon wir nun ausgehen – wir alle sie möglichst gut einschätzen können sollten. Meines Erachtens benötigen daher idealerweise alle Menschen zur medien- und gesellschaftskritischen Reflexion von KI Hintergrundwissen über ihre grundlegenden

Funktionsweisen sowie Strukturwissen zu spezifischen Stärken und Schwächen der Technologie sowie der auf ihr aufsetzenden konkreten Technik, wie Systemen und Tools.

Anknüpfend an die zuvor entfaltete Arbeitshypothese zu Bildung, die die kritisch-reflexive Auseinandersetzung des Individuums mit dessen Umwelt betont, habe ich im Vortrag mit ein paar visuellen Eindrücken einen exemplarischen ethnologisch-soziologischen Blick in unsere Umwelt geworfen. Die Visualisierungen sollten zum Nachdenken über verbreitete, aber vereinfachende, Annahmen anregen, die in den letzten Jahren die öffentlichen Diskurse dominierten: Ist es gefährlich, wenn Maschinen Fahrzeuge bewegen? Vereinsamen wir hinter den Bildschirmen und Displays? Wird Chaos herrschen, wenn Maschinen ‚entscheiden‘? Oder: Macht uns ‚Kolleg*in KI‘ arbeitslos?

Für diesen Text werfe ich – passend zum Titel der Ringvorlesung *Kollaboration mit Kollegin KI* – einen exemplarischen Blick in eine mögliche neue Arbeitswelt, in der die Arbeitsteilung zwischen natürlicher und künstlicher Intelligenz zur Normalität geworden ist. Ich habe freilich keine Glaskugel und selbstverständlich verfügt keine wissenschaftliche Disziplin über zweifelsfreie und etablierte Methoden für Aussagen über mögliche Zukünfte. Über den historischen Vergleich von *körperlicher* Arbeit durch menschliche Muskelkraft, die sich in den letzten Jahrhunderten aufgrund technischer Innovationen wie beispielsweise die der Dampfmaschine stark verändert hat, mit *geistiger* menschlicher Arbeit, die sich bereits durch Rechenmaschinen, wie den (Personal)Computer veränderte (vgl. u. a. Coy 1995; Keil 1994) und nun durch KI-Werkzeuge noch einmal deutlicher verändern wird, lassen sich aber zumindest ein paar konzeptionelle Aussagen treffen.

4.1 KI und (geistige) Arbeit

Seit der Erfindung, der Weiterentwicklung und der gesellschaftlich breiten Nutzung des Computers können Menschen nicht nur physische Arbeit an Maschinen delegieren, sondern auch geistige Arbeiten (vgl. u. a. Coy 1995; Keil 2003; Krämer 1988; Schelhowe 1997). Mit dem (Personal) Computer erhielt der Mensch ein vielseitiges *Werkzeug* für vielfältige geistige Aufgaben. Mit KI gerät die Werkzeugmetapher allerdings an ihre Grenzen, da das KI-Werkzeug nicht mehr handgeführt ‚be-dient‘ oder eigens programmiert werden muss (vgl. Knaus/Engel 2015, S. 27 f.; Knaus 2020b, S. 44 f.), sondern nun *wirklich* zum ‚Automaten‘ – im eigentlichen Wortsinne *αὐτός* (autos: selbst) und *μεν* (men: denkend bzw. wollend) – wurde (vgl. weiterführend auch Coy 1995). Obwohl die KI-Werkzeuge derzeit noch nicht sehr ausgereift sind, lässt sich diese These bereits durch genaue Beobachtungen bestätigen: Wie zuvor erwähnt, können Kreativarbeiter*innen wie Übersetzer*innen sich Entwürfe erstellen lassen und dadurch von Akkordarbeiter*innen zu Gutachter*innen werden (vgl. Siepmann 2023), Wissenschaftler*innen (vgl. DFG 2023) und Sachbearbeiter*innen werden von Routinearbeiten entlastet, Developer*innen und Programmierer*innen erhalten beim Coden, der ungeliebten Dokumentation oder der aufwändigen Fehlersuche Unterstützung – um nur einige wenige Möglichkeiten zu nennen (vgl. weiterführend Gmyrek/Berg/Bescond 2023 oder auch Knaus/Merz/Junge 2023). Eine noch breitere und umfänglichere Nutzungspraxis ist zu erwarten, wenn KI-Tools in Suchmaschinen (wie *Bing AI* oder *Google Bard*), weiteren Browsern (wie heute bereits in Safari, Chrome oder Opera) und in (weiteren) Applikation des Office-

Pakets aufgenommen werden oder das Programmieren zum Beispiel mittels *AlphaCode* von Alphabet (Dachorganisation des Tochterunternehmens Google LLC) mit ‚natursprachlichen‘ Anweisungen zur Selbstverständlichkeit wird.

Auch die wiederholte Aufbereitung (journalistischer) Texte durch automatisiertes Zusammenstellen von Nachrichten (zum Projekt ‚Genesis‘ vgl. Mullin/Grant 2023) oder die Mehrfach-Kompilation von Lifestyle-Informationen (zu *Burdas* 99-Pasta-Rezepten vgl. u. a. Ernst 2023) zeigen eindrücklich die Möglichkeiten, markieren gleichzeitig auch die Grenzen der Sinnhaftigkeit wiederholter Re-re-re-...Produktion von Informationen. Eine Übersetzung ist zwar streng genommen auch eine Re-Produktion einer Information, aber sie erschließt weitere Empfängergruppen. Bereits aktuell verfügbare Werkzeuge wie *Google Translator*, *DeepL* oder auch Textgeneratoren können beim Erstellen mehrsprachiger Texte oder Sprachübersetzungen eine große Hilfe sein. Nicht wenige Übersetzer*innen und Korrektor*innen befürchten daher, dass gerade die Weiterentwicklung von KI-Tools dazu führen könnte, dass ihr Tätigkeitsfeld auf längere Sicht verschwinden wird – vergleichbar wie der Beruf des Schriftsetzers nach der Erfindung des Desktoppublishing oder zahlreiche weitere Berufsbilder, die im Zuge von Digitalisierung und Automatisierung deutliche Veränderungen erfuhren oder ‚rationalisiert‘ wurden. Aus gesellschaftlicher Sicht (vgl. Gmyrek/Berg/Bescond 2023) sind diese Rationalisierungsprozesse verständlich, für das Individuum aber oft problematisch, da (Erwerbs-)Arbeit eine immense Bedeutung für uns Menschen hat. Für die Menschheit mag es eine Erleichterung sein, wenn zukünftig Roboter zum Beispiel selbständig Tauchgänge in Chemikalien durchführen und diese gefährliche Aufgabe nicht länger von Menschen erledigt werden muss, aber der qualifizierte Industrietaucher muss sich beruflich umorientieren und kann seinen Platz in der Gesellschaft verlieren. Gemäß der „Vermächtnis-Studie“ von infas schätzen 86 Prozent der Deutschen Erwerbsarbeit als *sehr wichtig* ein (vgl. infas 2018). Für die Mehrheit ist Erwerbsarbeit mehr als lediglich eine finanzielle Einnahmequelle, denn sie stellt überdies ein bedeutsames Identifikationsmerkmal für das Individuum innerhalb der Gesellschaft dar. Deutlich wird dieses Identifikationsmerkmal auch in verbreiteten Nachnamen, wie Becker, Bauer, Fischer, Schneider, Schlosser, Schreiner, Müller, Wagner oder Zimmermann, die auf die beruflichen Tätigkeiten der Vorfahren hinweisen.

Wahrscheinlicher als ein Verdrängen ganzer Berufszweige ist aber wohl eher, dass sich Berufsfelder *verändern* werden, wie die International Labour Organization – ILO in einer globalen Studie zeigt (vgl. Gmyrek/Berg/Bescond 2023). So zum Beispiel könnte sich das zuvor angesprochene Berufsbild des Übersetzers „vom Akkordarbeiter zum Gutachter“ verändern (Siepmann 2023, S. 494): Die Tätigkeit von Übersetzer*innen war bisher in hohem Maße durch die intensive (und oft auch mengenentlohnte) Textarbeit geprägt. Mithilfe von KI-basierten Übersetzungstools ersparen sich Übersetzer*innen das Übersetzen eines Texts von Grund auf und können sich – ähnlich wie Gutachter*innen – auf die kritische Prüfung, kulturelle Adaption und sprachliche Verbesserung des maschinell entworfenen Texts konzentrieren. Das Programmieren ähnelt dem Übersetzen sehr: Wer codet, übersetzt (menschliche) Anweisungen in eine maschinenverständliche Sprache. Verständlicher Weise kann diejenige, die diese Sprache besonders gut beherrscht – nämlich die Maschine selbst – diese Übersetzung auch selbst vornehmen: Textgenerierende KI-Tools wie *Copilot X* von GitHub, *Code Whisperer* von Amazon,

AlphaCode von Alphabet/Google, *Llama* von Meta können Programmcode besonders zuverlässig und schnell generieren oder auch Fehler im Programmcode aufspüren (Shani 2023) – eine Arbeit, die für Humanprogrammierer*innen in der Regel sehr (zeit-)aufwändig ist. Berufsbilder in der Softwareentwicklung können sich dadurch vom Fokus auf *Coding* hin zum *Developing* entwickeln: Statt stundenlangem Tippen und Fehlersuchen, können sich Entwickler*innen künftig stärker auf die kreative Seite ihres Jobs konzentrieren – der Konzeption, der Validierung und qualitativen Verbesserung von Applikationen sowie deren Usability (UI/UX). Mit dieser sehr erfreulich klingenden Arbeitserleichterung und Effizienzsteigerung sind jedoch zwei Herausforderungen für das Berufsbild verbunden: Da das zeitaufwändige manuelle Coden (beziehungsweise das Kopieren und Kompilieren von Code) von Humanprogrammierer*innen größtenteils wegfällt, steigen womöglich die Erwartungen an die Entwicklungsleistungen der Programmierer*innen, gleichzeitig wird ihre Aufgabe geistig anspruchsvoller. Spätestens an diesem Punkt gerät die aus der Zeit der Industrialisierung stammende Verknüpfung individueller Arbeitsleistung mit einer bestimmten Zeiteinheit (wie einer definierten täglichen Arbeitszeit) an ihre Grenzen. Mit der zu erwartenden Produktivitätssteigerung wird einhergehen, dass zur Umsetzung eines definierten Softwareprojekts weniger Arbeitsstunden beziehungsweise eine geringere Anzahl Developer*innen benötigt werden (vgl. Eloundou et al. 2023; Gmyrek/Berg/Bescond 2023) – etwa vergleichbar mit dem geringeren Bedarf an Arbeitskräften bei heutigen Bauprojekten in Relation zu Bauvorhaben in der Antike, in denen noch keine motorbetriebenen Baumaschinen den Großteil der harten körperlichen Arbeit abnehmen konnten. Die physische Unterstützung des Menschen durch (Bau-)Maschinen führte zwar zur Veränderung des Berufsbilds eines (Bau-)Arbeiters – plakativ ausgedrückt: vom Lastenträger zum Maschinenführer –, aber keineswegs zu struktureller Arbeitslosigkeit im Baugewerbe. Das Gegenteil war der Fall: Neue technische Möglichkeiten haben neue Bauformen und Baustile geprägt und die mit der Maschinennutzung einhergehende Produktivitätssteigerung ermöglichte Vorhaben, die alleine mit menschlicher Muskelkraft kaum zu stemmen beziehungsweise finanzierbar gewesen wären.

Die strukturellen Parallelen zwischen der Bedeutung der Industrialisierung für die Transformation physischer menschlicher Arbeit und der Bedeutung heutiger KI-Tools für die Veränderungen geistiger Arbeit ließen sich beliebig weiterführen. M. E. zeigt aber bereits dieser knappe Exkurs recht anschaulich, dass beide Transformationsprozesse grundsätzlich vergleichbar sind und da einer der beiden Prozesse – der der körperlichen Arbeit – bereits weit fortgeschritten und daher retrospektiv analysierbar ist, kann dieser möglicherweise Aufschluss über Möglichkeiten und Herausforderungen des aktuellen Transformationsprozesses geben.

Aus diesem Grund und da KI nicht die erste technische Innovation ist, die die geistige Arbeit von Menschen erleichtert, möchte ich zur Illustration die *strukturellen Parallelen* zwischen der Erleichterung körperlicher und geistiger menschlicher Arbeit durch Maschinennutzung noch etwas weiter ziehen und im folgenden – zugegeben etwas hemdsärmeligen, dafür aber leicht verständlichen – Beispiel vier unterschiedliche Stufen der Werkzeugnutzung unterscheiden (vgl. Tabelle 1): Wer schon einmal einen Graben oder eine Grube mit eigener Kraft ausgehoben hat, kann einschätzen, welche Arbeitserleichterung mit der Nutzung einer motorbetriebenen Baumaschine (wie zum Beispiel einem Bagger) verbunden ist. Dabei kann man vier

Modi der physischen Erdbewegung unterscheiden: Das Graben ohne jegliche Hilfsmittel mit bloßen Händen (1), das Graben mit geeigneten Handwerkzeugen (2), wie einem Spaten, die Erdbewegung mit einer mechanischen oder hydraulischen Baumaschine, die von einem Menschen geführt wird (3) oder mit einer autonomen Baumaschine ohne Fahrer*in (4). Diese einfache Unterscheidung ist auch hilfreich, um den Grad der Unterstützung von Werkzeugen oder Tools zu visualisieren, die die geistige Arbeit von Menschen unterstützen: Das Graben von Hand wäre mit der handschriftlichen Erstellung eines Texts (1) – eines *Manus Skripts* im wahrsten Sinne des Wortes – vergleichbar. Das Erstellen eines Textes mit einem Textverarbeitungsprogramm (2) wäre mit der Nutzung einer Schaufel oder eines Spatens (einem Handwerkzeug) vergleichbar; wenn die oder der Textproduzent*in eine Rechtschreibkorrektur oder einen Thesaurus nutzt (3), dann wäre dies mit dem Graben mit Unterstützung – wie mit einem motorbetriebenen Bagger – vergleichbar und die Nutzung eines KI-Tools wie *DeepL* zur Übersetzung oder eines KI-basierten LLM wie (Chat)GPT zur Texterstellung (4) wäre mit dem autonomen Erdaushub (wie im Tagebergbau) vergleichbar (vgl. zu vier Arten der Autor*innen-schaft weiterführend auch Bahjor 2022).

Tabelle 1: Visualisierung des Unterstützungsgrades physischer und geistiger Arbeit durch Werkzeuge/Tools und KI

physische Arbeit am Bsp. von Erdaushub	geistige Arbeit am Bsp. von Texterstellung
Erdaushub mit (bloßen) Händen	handschriftliche Texterstellung („Manuskript“)
Graben mit einem Handwerkzeug, wie einer Schaufel oder einem Spaten	Texterstellung mit einem Textverarbeitungsprogramm, wie Word, Pages oder Writer
Erdaushub mit einem menschengeführten, motorbetriebenen Bagger	Nutzung unterstützender Tools/Applikationen, wie einer Rechtschreibkorrektur oder Thesaurus
autonomer Erdaushub, wie im Tagebergbau	automatisierte Texterstellung oder Übersetzung von (Fremd-)Sprachen oder (Computer-)Code

Ob es nun um körperliche oder geistige Arbeit geht, Zwänge durch (anstrengende oder wiederkehrende) Beschäftigungen zu reduzieren und dadurch Freiräume zu gewinnen, scheint ein übergeordnetes Ziel menschlichen Innovationsbestrebens zu sein. Diese These ist in Anbetracht anstrengender körperlicher Arbeit naheliegend. Valide ist sie aber auch in Bezug auf bisherige Innovationen, die die geistige Arbeit des Menschen nachhaltig verändert haben – wie beispielsweise die Schrift und der Buchdruck, Rechenmaschinen wie der Taschenrechner, das WWW, die kollaborativ gepflegte Wikipedia oder auch Suchmaschinen. Wie andere Medien und Technologien zuvor, haben diese Innovationen unsere menschlichen Sinne und Möglichkeiten erweitert (vgl. weiterführend zu *Extensions of Man* McLuhan 1964, S. 7 f.). Basierend auf den Erfahrungen mit diesen Innovationen, die in den vergangenen Jahrzehnten sowohl physische als auch geistige Arbeit transformiert haben, können wir rückblickend auch für gegenwärtige Herausforderungen um KI-Systeme wertvolle Erkenntnisse gewinnen. Denn keine mediale oder technologische Neuerung besaß lediglich Potentiale, sondern brachte stets auch

Herausforderungen mit sich. Auch die geistige und kreative Zusammenarbeit von Mensch und Maschine birgt verständlicherweise zahlreiche Fragen und Herausforderungen sowie Normierungs- und Organisationsbedarfe.

An dieser Stelle möchte ich mich mit der stichwortartigen und exemplarischen Aufzählung der mit der Übernahme geistiger Arbeit durch KI verbundenen Schwierigkeiten und Herausforderungen begnügen, da sie größtenteils im weiteren Verlauf dieses Beitrags noch ausführlicher²³ besprochen werden: Umgestaltung von Berufsbildern und Arbeitszusammenhängen oder sogar individueller Verlust des Arbeitsplatzes, höhere Abhängigkeiten von Technik, Auswahl und Qualität der (umfanglichen) Datenbestände beziehungsweise Schwierigkeiten beim Lernen mit kleinen Datensätzen, fehlende Transparenz bezüglich der Herkunft der Daten und der Genese der Texte beziehungsweise der Antworten (vgl. weiterführend zu Explainable AI, Abschnitt 5.2), Fehler wie sogenannte ‚Halluzinationen‘ (ausführlicher im folgenden Abschnitt), rechtliche Fragen, wie Schadensersatzforderungen für Nutzung von Content für Trainingszwecke oder urheberrechtliche und datenschutzrechtliche Fragen, Beeinflussung der Meinungsbildung durch so genannte ‚Echokammern‘ oder ‚Filterblasen‘ (vgl. auch Fußnote 22).

Da generative KI stets nur re-produziert, d. h. lediglich das nachahmt, was sie aus den (mehr oder weniger umfanglichen) Trainingsmaterialien zur Parametrisierung kennt, reproduziert sie auch die in den Trainingsdaten enthaltenen Wertungen, Vorstellungen und Stereotype (vgl. u. a. Beranek 2020, S. 76 f.; Knaus 2020b, S. 40 ff.): Eine KI kann nicht out-of-the-box ‚denken‘, das macht sie nicht selten konservativer als manche dörfliche Stammtischrunde. Aber – anders als in der Stammtischrunde – bleiben die Stereotype, die eine KI transportiert, re-kontextualisiert und re-produziert oft verborgen und entsprechend unreflektiert. So würde beispielsweise die Bitte an eine bildgenerierende KI ein „fotorealistisches Bild eines Flugbegleiters“ – hier absichtlich in der männlichen Form – zu erstellen, das Bild einer *weiblichen* Flugbegleiterin erzeugen, da in der Datengrundlage nicht nach Geschlecht unterschieden wird. Die Bitte um eine „fotorealistische Büroszene im Großraumbüro“ würde typischerweise ein Bild generieren, auf dem die männliche Person stehend etwas erklärend oder beauftragend zu sehen ist, während die weibliche Person anerkennend zuschaut oder konzentriert am Telefon oder Computer arbeitet. Auch die Titelfolie meines Vortrags verrät viel darüber, was die KI sich unter einer ‚Collegerunde‘ vorstellt: Die bildgenerierende KI *Midjourney* wählte im Kontext des Promptbestandteils „at college“ jeweils zielsicher das Vorlesungsformat und keine Seminarform (obwohl die Seminarform im angloamerikanischen Hochschulkontext viel üblicher ist) und ließ sich davon auch durch das Weglassen und sogar das Ersetzen der Promptbestandteile „presentation“ und „auditorium“ partout nicht davon abbringen.

²³ Noch ausführlicher werden wir uns in der 24. Ausgabe der LBzM (2024) mit der gesellschaftlichen und kulturellen Bedeutung von Künstlicher Intelligenz und dem humanen und konstruktiv-kreativen Miteinander von Mensch und Maschine befassen (vgl. Knaus/Merz/Junge 2023).



Abbildung 1: Titelfolie des Vortrags „KI in Schule und Hochschule – Impulse aus bildungstheoretischer Perspektive“, erstellt mittels des bildgenerierenden KI-Tools *Midjourney* [Prompt: realistic happy artificial intelligence (AI), who looks like a robot, delivering a presentation to students at college, the auditorium is bright, students are happy]

Diese Beispiele zeigen recht anschaulich, dass bildgenerierende KI-Tools in das Trainingsmaterial ‚eingeschriebene‘ (präziser: parametrisierte) Werte und Vorstellungen mehr oder weniger transparent re-produzieren. Dies gilt freilich auch für textgenerierende KI, in der bestimmte Begriffsverwendungen oder textliche Nuancen die eingeschriebenen Stereotype verraten können. Hieraus wird deutlich, dass zur Erreichung des Bildungsziels des kritischen und selbstbestimmten Umgangs mit Medien und Technik Gelegenheiten zur Reflexion der Hintergründe und Funktionsweisen sowie ihrer strukturellen Stärken und Schwächen geschaffen werden müssen. Gleichmaßen wird hierin auch ein interessantes reflexives Potential bildgenerierender KI deutlich, wenn per Text-to-Image-Generatoren Stereotype visualisiert und dadurch aufgedeckt werden können. Wie so oft in der Geschichte der Medieninnovationen entsteht durch die jeweilige Entwicklung und gesellschaftliche Nutzung nicht nur ein Bildungsbedarf, sondern die Lösung zur bildungsbezogenen Reflexion wird durch die Technik oder das Medium gleich ‚mitgeliefert‘.

4.2 KI und Informationsrecherche

Recherchen – zum Beispiel von Studierenden und Berufseinsteiger*innen (vgl. Nagel/Zlatkin-Troitschanskaia/Fischer 2023) – beginnen heute in der Regel online mit einer Websuche. Suchmaschinen liefern einen schnellen Zugang zu unkommentierten Datensätzen in Form von oft seitenlangen Linksammlungen. Mit der Onlinerecherche waren bisher primär drei Anforderungen verbunden: Da die Qualität des Suchergebnisses abhängig von der Passung der Suchbegriffe sowie ihrer Kombination (ggf. auch mittels boolescher Operatoren) ist, wurde es im hohen Maße durch das Wissen über Suchstrategien und das fachliche Vorwissen des Recherchierenden bestimmt. Die Ergebnisliste konnte nur im Gesamt durchgesehen und nicht

durch präzisierende oder vertiefende Nachfragen angepasst und verfeinert werden. Außerdem musste jedes der Fundstücke vom Recherchierenden einzeln interpretiert und eingeordnet werden (ausgenommen bei der Verwendung sogenannter Metasuchmaschinen, wie *MetaGer*)²⁴. Die Integration von Dialogsystemen (ChatBots) wie *Bing AI* oder *Google Bard* in (Web-)Suchmaschinen ermöglicht künftig Nachfragen zur Erweiterung oder Eingrenzung beziehungsweise Präzisierung des Suchergebnisses. Auch könnten Recherchierende durch KI-gestützte Analysen Informationen in Form von aufbereiteten Zusammenfassungen oder visualisierten Strukturmodellen erhalten.²⁵ Diese aufbereiteten Meta-Informationen könnten sogar durch erklärende Hinweise ergänzt werden, um zum Beispiel Kontexte besser einschätzen zu können oder Fachfremden das Verständnis von Fachtermini zu erleichtern. Ermöglicht wird hierdurch eine (noch) effizientere Informationsbeschaffung, die gleichermaßen auch tiefere Einblicke in komplexe Themen ermöglicht.

Diese vielversprechenden Potentiale und nicht zuletzt auch die Effizienzversprechen klingen sehr reizvoll – gerade für gestresste Wissenschaftende oder für Schüler*innen und Studierende, denen nur noch wenig Zeit bis zum Ablauf einer Deadline bleibt. Doch sind nicht genau die kluge Wahl und Kombination der Suchbegriffe sowie die Interpretation der Ergebnisse bereits wesentlicher Teil des Wissensproduktion? Und wer kennt nicht die Situationen, in denen sich bei der Durchsicht langer Linklisten oder auch nach der längeren Verfolgung eines ursprünglich wenig aussichtslosen Links plötzlich neue Erkenntnisse zur aktuellen Fragestellung auftun? Oder wenn man bei der Suche in der Bibliothek rechts und links im Magazin schaut und so das Buch findet, das die weitere Arbeit entscheidend prägt. Ein Outsourcen dieser Rechercheprozesse macht die Informationssuche effizienter, aber macht sie sie auch *effektiver*? Und woran erkennen wir, dass wir ein Thema wirklich durchdrungen haben oder alle Antwortoptionen auf eine Frage wahrgenommen und abgewogen haben?

Diese und weitere Herausforderungen sind mit der KI-gestützten Informationsrecherche verbunden. Eine davon nenne ich die (umfänglichere) *Uniformierung von Wissen*. Gemeint ist damit, dass wir nur noch das finden, was die (Such-)Maschinen bereits kennen oder als relevant erachten. Relevanz wurde schon bei klassischen Web-Suchmaschinen über die Linkpopularität (vgl. zum *PageRank*-Algorithmus auch Langville/Meyer 2006) sowie die Anzahl der Klicks, Verweildauer und Absprungraten anderer Nutzer*innen bestimmt: (Web-)Seiten und die auf ihnen enthaltenen Informationen sind für den Suchalgorithmus u. a. dann besonders relevant, wenn zahlreiche Links auf sie verlinken und möglichst viele andere Nutzer*innen die betreffende Seite aufrufen sowie innerhalb der Seite weiterklicken und möglichst lange verbleiben.

²⁴ Weiterführende Hinweise zu Web- und Meta-Suchmaschinen sowie Recherchetipps speziell für die wissenschaftliche Recherche von Studierenden finden sich in den beiden Kurzvideos „Wie findet man geeignete Literatur?“ (youtu.be/qjH66UJhb4E) und „Wie findet man geeignete Literatur – online... fast“ (youtu.be/PixHs5WysLk) der Videoserie *#Medienpädagogik im Schaukelstuhl* sowie ihren ausführlichen Beschreibungstexten.

²⁵ Speziell für die wissenschaftliche Recherche können KI-Tools wie *Elicit*, *ResearchBuddy* und *ResearchRabbit* unterstützen, indem sie Literaturübersichten generieren oder Verbindungen zwischen unterschiedlichen Studien und Arbeiten visualisieren.

Was kann es nun für den gesellschaftlichen Wissensbestand bedeuten, wenn KI-basierte Dialogsysteme, die stets nur das fragen können, was sie selbst kennen (vgl. Abschnitt 2), Menschen bei ihrer Informationsrecherche lenken und die Suchenden daraufhin primär auf die bereits hoch ‚gerankten‘ Ergebnisse klicken? Möglicherweise werden durch diese Praxis – wie bereits durch die Empfehlungsalgorithmen zur Korrektur „Meinten Sie...“ oder das automatisierte Vervollständigen bei der Eingabe von Suchbegriffen – als ‚relevant‘ gerankte Informationen zunehmend relevanter, während Inhalte weniger häufig verlinkter und besuchter Seiten sukzessive aus dem gesellschaftlichen Wissensbestand verschwinden. Die Informations- und auch die Meinungsvielfalt des Netzes können durch diese Praxen, die eigentlich eine ‚effiziente‘ Suche ermöglichen sollen, eingengt und entsprechend ‚universalisiert‘ werden.

Aufgrund dieser Re-Produktion des Populären und Standardisierten birgt die Verwendung von KI-basierten Suchalgorithmen bei der Informationsrecherche auch das im vorherigen Absatz bereits beschriebene Risiko der Verfestigung von Wertungen, Vorstellungen und Stereotypen: Offensichtliche oder oft auch nicht sichtbare in Texten und Bildern eingeschriebene Vorurteile und Stereotype können durch die unreflektierte Re-produktion generativer KI oder KI-basierter Dialogsysteme verstärkt werden, was zu einer unzureichenden oder einer verzerrten Sicht auf die Realität führen kann (vgl. weiterführend Beranek 2020, S. 76 f.; Knaus 2009, S. 210 ff.; Knaus 2020b, S. 40 ff.). Wie bereits angemerkt, kann aber beispielsweise bildgenerierende KI auch dabei helfen, diese üblicherweise versteckten Stereotype zu visualisieren und dadurch sicht- und reflektierbar zu machen.

Weitere Probleme sind denkbar, wenn textgenerierende KI-Tools dazu missbraucht werden, Social-Media-Plattformen mit (gezielter) *Desinformation* (oft auch als so genannte ‚Fake News‘ bezeichnet) zu ‚fluten‘. Das zuvor bereits beschriebene Relevanzkriterium der Popularität kann auf diese Weise leicht zur Meinungsbeeinflussung missbraucht werden. Besonders herausfordernd wird der Umgang mit Desinformationen, wenn überdies gefälschte Bilder einer bildgenerierenden KI oder (DeepFake-)Videos verwendet werden, die in der Öffentlichkeit als authentisch wahrgenommen werden – der Fake, der den *Papst im Daunenmantel* zeigt (vgl. Weiß 2023), ist hier zwar eines der bekanntesten, aber auch eines der harmlosesten Beispiele. Bisher haben wir es also eher noch mit ‚cheap fakes‘ als mit *deep fakes* zu tun, aber das ist sehr wahrscheinlich nur eine Frage der Zeit (vgl. Weck 2023).

Da wir Bilder aus der Natur kennen und – im Gegensatz zu Texten – den kritischen Umgang mit Bildern oder Bewegtbildern (bisher) kaum gelernt haben, stellen wir ihre Glaubwürdigkeit selten in Frage. Visuelle Codes im Allgemeinen sowie auch die konnotative Funktion von Bildern und Videos verfügen damit über einen direkten, unvermittelten Zugang zur menschlichen Wahrnehmung (vgl. Bruce 2008; Knaus 2009, S. 27–31; Knaus 2018, S. 95 f.; Knaus 2020a, S. 9).

Apropos Wahrnehmung: *Halluzinationen* sind Wahrnehmungen von etwas, das es nicht gibt. Wenn Menschen etwas wahrnehmen wofür in der physischen Welt keine nachweisbare Reizgrundlage existiert, dann sprechen Neurolog*innen von Halluzinationen. Als so genannte ‚KI-

Halluzinationen²⁶ werden Ergebnisse einer KI bezeichnet, die keinen Sinn ergeben, frei erfunden sind oder die mit den Trainingsdaten nicht übereinstimmen: Eine textgenerierende KI kann Falschmeldungen generieren, stellt hierfür Zusammenhänge falsch dar, schildert Sachverhalte verzerrt und kann hierfür sogar falsche oder nicht existierende Belege ‚erfinden‘. ‚KI-Halluzinationen‘ ergeben sich aus der spezifischen Funktionalität von Textgeneratoren, sie sind ein strukturelles Problem, das sich aus den statistischen Berechnungen ergibt: Denn wenn Textgeneratoren nach wiederkehrenden Mustern das jeweils nächste passende Wort vorhersagen (vgl. zur Funktionsweise von Textgeneratoren auch Abschnitt 2), dann führt das zwar zu beeindruckenden Texten, aber eben auch mit einer gewissen Regelmäßigkeit zu Aussagen, die nicht korrekt sein müssen.

Ein inzwischen recht bekanntes Beispiel ist der Fall eines Rechtsanwalts aus den USA, der ChatGPT offensichtlich für eine Suchmaschine hielt, sie nach Präzedenzfällen befragte und die erhaltenen Antworten ungeprüft übernahm (vgl. Wittenhorst 2023). Der Textgenerator erzeugte (präziser: fantasierte) aus seinen Trainingsdaten ganze Fallakten zusammen, die bis in „Details hinein einen Anschein von Echtheit verliehen – inklusive der üblichen gelehrigen Erläuterungen“, die der Anwalt nicht auf ihre Echtheit hin überprüfte. Wer das US-amerikanische Rechtssystem kennt, kann gut nachvollziehen, warum der Anwalt sich von einer KI helfen lassen wollte: Denn nach der Rechtsordnung in den USA – dem Fallrecht (*case law*) – stützt sich die Rechtsfindung primär induktiv auf vorangegangene Rechtsprechung zu vergleichbaren Fällen, auf so genannte Präzedenzfälle. Das Aufspüren dieser Präzedenzfälle, die bis in die Anfänge der Vereinigten Staaten zurückreichen können, ist nachvollziehbarerweise sehr aufwändig. Der Anwalt wollte sich also viel Arbeit beziehungsweise zahlreiche Arbeitsstunden für Rechtsassessor*innen sparen. Warum er jedoch kein einziges Aktenzeichen kontrolliert hat und damit die Standards und Methoden des eigenen Fachs vollständig ignorierte, ist kaum nachvollziehbar.

Damit komme ich auf eine der verlockendsten Angebote im Kontext der alltäglichen KI-Nutzung zu sprechen, der Möglichkeit, das *Nachdenken* gänzlich an eine KI abzugeben – und damit zugleich auf eine der großen Herausforderungen für den Bildungsbereich: Was antworten wir Schüler*innen oder Studierenden, die nichts mehr lernen wollen und stattdessen lieber die Möglichkeiten nutzen möchten, die ihnen selbstverständlich zur Verfügung stehen: Warum soll ich rechnen lernen, wenn es ein Taschenrechner schneller kann? Warum soll ich etwas (auswendig) lernen, wenn ich es jederzeit in der *Wikipedia* nachschlagen kann oder mit *Google* finde? Warum soll ich eine Fremdsprache lernen, wenn ich auch eine App auf meinem Smartphone nutzen kann? Und warum sollte ich das Formulieren von Texten und Orthografie lernen, wenn mir diese anstrengende Arbeit auch ein Textgenerator abnehmen könnte?

²⁶ Die Bezeichnung dieses strukturellen Problems von KI-Systemen als ‚Halluzination‘ ist irreführend (und könnte wohl auch selbst das Ergebnis einer Halluzination sein), der unscharfe und irreführende Begriff etablierte sich aber leider bereits binnen sehr kurzer Zeit.

4.3 KI in der Schule

Bereits aus den beiden vorherigen Unterkapiteln, in denen ich beschrieben habe, wie KI-Werkzeuge unsere Informationsrecherche und etablierte Prozesse der Wissensproduktion (vgl. Abschnitt 4.2) sowie geistige und kreative Arbeitsprozesse (vgl. Abschnitt 4.1) berühren und verändern ergibt sich bereits der Bedarf, dass Schule und Hochschule eine um digitalisierungsbezogene Aspekte erweiterte Medienbildung in ihre Curricula aufnehmen. Aus medienpädagogischer Sicht genügt jedoch der Fokus auf wünschenswerte Fähigkeiten für die ‚digitale (Berufs-)Welt‘ (vgl. u. a. KMK 2016; Vuorikari/Kluzer/Punie 2022) oder zur souveränen Bewältigung des Alltags nicht (vgl. u. a. Kommer et al. 2016). Die um technische und informatische Inhalte *erweiterte* Medienbildung bereitet aber nicht nur adäquat auf eine durch Mediatisierung und Digitalisierung geprägte Welt vor (vgl. u. a. Knaus 2018, S. 99–103; Knaus 2020a, S. 13); mit ihr wird außerdem auch die nötige Grundlage für medien- und gesellschaftskritische Reflexionsprozesse hinsichtlich technologischer und technischer Hintergründe (wie Funktionsweisen von KI) gelegt, die wiederum wesentliche Voraussetzungen für die kritische und konstruktive Auseinandersetzung mit den konkreten Möglichkeiten und Herausforderungen der auf ihr aufsetzenden Tools und Medien sind (vgl. Knaus/Schmidt/Merz 2023).

Wie bei zahlreichen Medieninnovationen zuvor ergibt sich aus der gesellschaftlichen Nutzung der Werkzeuge und Medien in Alltag und Beruf für Bildungseinrichtungen nicht nur eine Erweiterung ihres Bildungsauftrags, sondern mit Medieninnovationen waren stets auch Potentiale für Lehr-, Lern- und Bildungskontexte verbunden (vgl. u. a. Knaus/Engel 2018). Entsprechend sind auch mit KI-Werkzeugen und deren gesellschaftlicher Nutzung erneut Potentiale und Herausforderungen verbunden, die ich im Folgenden exemplarisch und knapp gegenüberstellen möchte.

Ein vielversprechendes Potential von KI im Bildungskontext liegt in der Nutzung von Chatbots als geduldige ‚Sparringspartner‘ beim Lernen: Die Chatbots können als *Assistent*innen* Schüler*innen zum Beispiel bei Hausaufgaben unterstützen oder personalisierte ‚Nachhilfe‘ anbieten. Durch ihre hohe Anpassungsfähigkeit und kontinuierliche Verfügbarkeit können sie den Lernprozess individuell begleiten und dadurch Eltern sowie Lehrende unterstützen und entlasten. Nicht zuletzt kann sich die KI-basierte Hausaufgabenbegleitung oder Nachhilfe auf das individuelle Lerntempo der einzelnen Schüler*innen einstellen und über einen längeren Zeitraum der individuellen Begleitung auch die Stärken und Schwächen der oder des Lernenden zunehmend besser einschätzen und gezielt fördern (vgl. hierzu auch ältere konzeptionelle Arbeiten zu *Virtual Learning Environments – VLE* oder *Personal Learning Environments – PLE* u. a. von Attwell 2007).

Ein weiteres vielversprechendes Potential der KI für die Schule liegt im Umgang mit Heterogenität und unterschiedlichen Leistungsstufen im Klassenverbund: Textgenerierende KI-Systeme können Lehrer*innen im Unterricht sowie bei der Unterrichtsvorbereitung dabei unterstützen, Lehrmaterialien umfangreicher zu individualisieren und dadurch Schüler*innen mit unterschiedlichen Leistungsständen leichter und gezielter gerecht zu werden. Dies ermöglicht eine umfassendere Individualisierung des Unterrichts und kann dadurch zur besseren Einbindung und gezielten Förderung *aller* Schüler*innen beitragen. Diese Unterstützung von Lehrenden mit dem Ziel der individuellen Förderung reicht von der individualisierten Vergabe von

Hausaufgaben sowie deren Vorkorrektur bis zur Auswertung von Lernprozessen mit dem Ziel der Kompetenzprüfung und -entwicklung sowie dem Entwurf persönlicher Förderpläne auf Basis der gesammelten Auswertungen (vgl. zu *Learning Analytics* bzw. *Classroom Analytics* u. a. Hartong 2019; Reinholz/Stone-Johnstone/Shah 2020).

Wie in den vorherigen Ausführungen bereits deutlich wurde, sollte auch bei der Verwendung von KI zur Einschätzung von Lernfortschritten und dem Entwickeln von Förderplänen den nutzenden Lehrenden bewusst sein, dass die KI ihnen Empfehlungen ausspricht, um ihre Expertise als Pädagog*innen zu stützen – etwa so, wie ein Handwerkzeug die menschliche Muskelkraft gezielt unterstützen kann, sie aber keineswegs ersetzt (vgl. Abschnitt 4.1). Hier unterscheidet sich die pädagogische Praxis verständlicherweise nicht von anderen Berufsfeldern, wie der Medizin oder der Justiz, in denen ebenfalls KI-basierte Expertensysteme stärker zum Einsatz kommen sollen. Als eine Selbstverständlichkeit beim Einsatz KI-basierter Unterstützungssysteme im schulischen Kontext sollte gelten, dass weder die Autonomie der Lernenden beeinträchtigt noch ihre Privatsphäre verletzt wird (vgl. Deutscher Ethikrat 2023, S. 248 f.). Mit diesen – in pädagogischen Zusammenhängen eigentlich selbstverständlichen – Hinweisen wird bereits deutlich, dass der Einsatz KI-basierter Systeme in Schule und Unterricht sowie in der Nachhilfe oder Hausaufgabenbetreuung verständlicherweise nicht nur Potentiale, sondern auch einige Herausforderungen birgt und mit ihm auch nicht wenige Befürchtungen verbunden sind.

Die wohl weitreichendste Befürchtung zeigt sich in der dystopischen Vision einer (umfanglichen oder sogar gänzlichen) Ersetzung von Lehrpersonen durch Maschinen. Diese Dystopie ist keine Befürchtung angehender Lehrer*innen, sondern wird sogar in der Stellungnahme des Deutschen Ethikrats *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz* antizipiert und mit dem aktuellen Personalmangel im Bildungsbereich in Verbindung gebracht: „Eine vollständige Ersetzung von Lehrkräften [...] ist auch nicht dadurch zu rechtfertigen, dass schon heute in bestimmten Bereichen ein akuter Personalmangel und eine schlechte (Aus-)Bildungssituation herrschen“ (Deutscher Ethikrat 2023, S. 249). Interessant ist, dass sich diese Dystopie wie ein ‚roter Faden‘ durch die Mediengeschichte zieht: Sobald technische Innovationen über Potentiale für Lehr-, Lern- oder Bildungskontexte verfügen, wird über eine (prospektive) Ersetzung von Lehrer*innen gesprochen (vgl. Meyer-Drawe 1995; Schorb 1994) und als Gegenreaktion die Relevanz von Lehrpersonen in der Schule oder der Lernförderlichkeit menschlicher Beziehungen unterstrichen (vgl. u. a. Felten 2010; Lankau 2022). Schon Bücher und Zeitungen standen seinerzeit im Verdacht, das bis dahin herrschende ‚Informationsmonopol des Lehrers‘ zu gefährden oder Lehrenden sogar gänzlich ihre Arbeitsgrundlage zu entziehen. In wiederkehrenden Wellen wechselten sich Medieninnovationen mit Potentialen für Lehr- und Lehrkontexte sowie unrealistischen Zukunftsvisionen und Widerständen gegen die damit einhergehenden Veränderungen ab (vgl. u. a. Knaus/Engel 2016). Die Liste der Beispiele, die euphorisch als ‚Bildungsrevolutionen‘ angekündigt wurden und in der Kritik standen, eines Tages Lehrer*innen ersetzen zu können, ist lang: Schulfilm und Schulfernsehen (vgl. Cuban 1986; Schorb 1982; Schorb 1994), programmiertes Lernen beziehungsweise programmierter Unterricht und Sprachlabore (vgl. u. a. Skinner 1954; Freudenstein 1975), ‚Lehrmaschinen‘ (vgl. Ferster 2014) oder auch VLE/PLE und MOOCs im Hochschulkontext (vgl. u. a. Attwell 2007; Baggaley 2013; Mayrberger 2010, S. 288–296).

Die Geschichte lehrt, dass diese Befürchtungen nicht eintraten (vgl. u. a. Meister 2008, S. 514 f.; Schorb 1994, S. 204) und daher sehr wahrscheinlich auch im aktuellen Kontext nicht eintreten werden – obschon die (Interaktions-)Möglichkeiten der Technik diesmal wirklich an historische Vorstellungen einer ‚Lehrmaschine‘ heranreichen (vgl. u. a. Skinner 1954 und kritisch dazu u. a. Meyer-Drawe 1995, S. 363). Selbst Maschinen mit umfänglichen und elaboreierten Interaktionsmöglichkeiten werden keine Lehrer*innen ersetzen, da ein vollständiger Ersatz nicht sinnvoll ist – zumal in pädagogischen Kontexten die Beziehungsarbeit im Fokus steht (vgl. u. a. Deutscher Ethikrat 2023, S. 248; Hartong 2019) und der Schule nicht nur die Funktion der (beruflichen) Qualifizierung zukommt, sondern auch die der Persönlichkeitsbildung, der Enkulturation und der gesellschaftlichen Integration der Schüler*innen (vgl. u. a. Fend 1979). Aus diesem Grund darf die Beziehungsarbeit der Lehrenden, aber auch das soziale Miteinander der Schüler*innen im Klassenverbund nicht zu kurz kommen (vgl. Deutscher Ethikrat 2023, S. 182). Ist es nicht merkwürdig, dass sich diese Befürchtung dennoch so konsequent durch die Mediengeschichte zieht? Es kommt doch auch heute niemand auf die Idee, dass Ärzte ihre Beschäftigungsgrundlage verlieren, weil Patient*innen nach ihren Symptomen ‚googlen‘ oder KI-basierte Expertensysteme medizinische Diagnosen vornehmen – und darin übrigens sogar sehr gut sind (Singhal/Azizi/Tu et al. 2023)? Expertensysteme übernehmen keine Verantwortung für medizinische Diagnosen und elektrische Handwerkszeuge machen auch keine Installateur*innen arbeitslos – Werkzeuge sind nun mal Werkzeuge und können Menschen gezielt unterstützen, aber sie sind keine Automaten (vgl. Abschnitt 4.1; weiterführend Knaus/Engel 2015).

Ein weitere Herausforderung liegt in der zuverlässigen und fairen Kompetenz- und Leistungsüberprüfung: Traditionelle Bewertungsmethoden könnten einerseits durch KI-Systeme erweitert werden und dadurch Lehrende in ihrer aufwändigen Korrekturarbeit unterstützen – zum Beispiel durch Korrektur- und Bewertungsempfehlungen, die nicht zuletzt auch Möglichkeiten einer individualisierten Bewertung und ausführlicheren persönlicheren Rückmeldung zu Essays oder Fehlern in Klausuren ermöglichen könnten. Andererseits steigen aber auch die Aufwände, überwachte Prüfungssituationen herzustellen, da die KI-Tools von Schüler*innen auch zur Täuschung und zum Erschummeln guter Bewertungen missbraucht werden können.

Eine weitere – bereits kurz angesprochene – Sorge ist, dass Schüler*innen Aufgaben ausschließlich der Maschine überlassen. KI-Tools können zwar eine nützliche Unterstützung sein, sollten jedoch bestimmte (Übungs-)Aufgaben für Schüler*innen nicht übernehmen oder gar als Ersatz für eigenständiges kreatives Schreiben beziehungsweise Gestalten, kritisches Denken und Problemlösen missverstanden werden – zumal diese Prozesse miteinander verbunden sind (vgl. u. a. Bereiter 1980). Nicht zuletzt zeigt sich hierin erneut die Relevanz einer gut (aus-)gebildeten Lehrperson, die adäquat abwägen und entscheiden kann, welche Übung für ein tieferes Verständnis der jeweiligen Materie nötig ist und welche Aufgaben getrost eine KI übernehmen könnte. Immerhin haben Lehrer*innen sehr ähnliche Herausforderungen bereits bei vorangegangenen Medieninnovationen gemeistert – just das sollte dieser Abschnitt illustrieren und damit einen Beitrag zur differenzierten kritisch-konstruktiven Diskussion des Verhältnisses von KI und Schule leisten.

4.4 KI in der Hochschule und der Universität

Auch im universitären Kontext ergeben sich vielfältige Potentiale für die sinnvolle Nutzung von KI-basierten Tools in Forschung und Lehre. Eine vielversprechende Möglichkeit liegt – analog zur Schule und der häuslichen Nachhilfe – in der individualisierten Unterstützung des *Selbststudiums*: Wie auch zuvor für die schulische Bildung beschrieben, können KI-Systeme universitäre Lehrmaterialien und Lernumgebungen so anpassen, dass sie den unterschiedlichen Voraussetzungen der Studierenden gerecht werden. So können KI-Tools Lehrende bei der Erstellung maßgeschneiderter Lehrmaterialien und Lernumgebungen unterstützen, in denen die individuellen Bedarfe von Studierenden unterschiedlicher Fächer und Studienfortschritte, fachbezogene Wissensstände und Lerngeschwindigkeiten, Motivationen, Herkünfte und Sprachen sowie die persönlichen Lerninteressen und Forschungsschwerpunkte zum Tragen kommen. Diese personalisierten Lehrmaterialien und Lernumgebungen würden es den Studierenden erlauben, ihre Lernprozesse eigenständig zu gestalten und zu steuern (vgl. auch hierzu ältere konzeptionelle Arbeiten zu *VLE* oder *PLE* u. a. von Attwell 2007; Mayrberger 2010, S. 288–296). Dies ermöglicht allen Studierenden eine umfassendere Individualisierung des Lernens sowie Fortschritte im eigenen Tempo. Individuelles Lernen erfordert von den Studierenden eine erhöhte Selbstverantwortung und ist auch ein wichtiger Beitrag zur Bildungsgerechtigkeit: Studierende mit unterschiedlichen Voraussetzungen und Hintergründen erhalten die Möglichkeit, auf ihre eigene Weise zu lernen (vgl. weiterführend auch Rath/Maisenhölder 2021). Nicht zuletzt lässt sich aus lerntheoretischer Sicht argumentieren, dass jegliches Lernen selbstgesteuerte, autopoietische Erfahrungskonstruktion und damit stets abhängig von bereits erlernter und gelebter Wirklichkeit ist (vgl. weiterführend u. a. Reich 2008; Siebert 2005, S. 33; Knaus 2013, S. 25) – aus lerntheoretischer Sicht ist eine individuell gesteuerte Unterstützung und Begleitung ideal (vgl. u. a. Mayrberger 2010; Rath/Maisenhölder 2021), die mit KI-basierten ChatBots nun real werden kann. Denn wie auch in der Schule können ChatBots als ‚Sparringspartner*innen‘ auch Studierende im Selbststudium unterstützen, indem sie Fragen beantworten oder auch mit klugen Fragen zum kontroversen Austausch anregen (vgl. auch Abschnitt 5.4) und damit den Studierenden helfen, ihr kritisches Denken und ihre Problemlösungsfähigkeiten zu schärfen. Ein inspirierendes Praxisbeispiel, das diese Potentiale im Hochschulkontext veranschaulicht, ist das Projekt *VoluProf*. Dieses in Zusammenarbeit mit dem Deutschen Forschungszentrum für Künstliche Intelligenz – DFKI, der Universität Rostock und dem Fraunhofer-Institut initiierte Projekt zeigt, wie Studierende den Vorlesungsinhalt individuell anhalten und ihrer beziehungsweise ihrem Professor*in als „animierten 3D-Avatar in fotorealistischer Qualität“ trotz des Großgruppenformats individuelle Fragen stellen können. Mit diesem Projekt möchten die Projektpartner verdeutlichen, wie das herkömmliche Vorlesungsformat KI-gestützt bereichert werden und das „Lern-Erlebnis immersiver und intensiver“ gestaltet werden kann (vgl. [fraunhofer.de/de/presse/presseinformationen/2023/august-2023/voluprof-ermoeglicht-individuelle-und-interaktive-online-vorlesung.html](https://www.fraunhofer.de/de/presse/presseinformationen/2023/august-2023/voluprof-ermoeglicht-individuelle-und-interaktive-online-vorlesung.html)).

Insbesondere im Hochschulkontext ist es von großer Bedeutung sicherzustellen, dass Studierende nicht das eigenständige Denken und die kritische Analyse einer Maschine überlassen. Die Unterstützung durch KI sollte deshalb gerade im Studium als sinnvolle Ergänzung und nicht als Ersatz für menschliches Denken betrachtet werden – zumal im Hinblick auf die hier im Zentrum stehenden Textgeneratoren der Wert des kreativen Scheibens als Denkwerkzeug

nicht unterschätzt werden sollte (vgl. u. a. Bahjor 2022; Bereiter 1980).

Die Integration von KI in die Hochschulbildung wirft auch Fragen zur Bewertung und der zuverlässigen und fairen Kompetenz- und Leistungsüberprüfung auf (vgl. u. a. Marchand 2023; Salden/Lordick/Wiethoff 2023; Weißels 2021). Insbesondere tradierte Prüfungsformen wie schriftliche (Haus- und Qualifikations-)Arbeiten und (OpenBook- oder Online-)Klausuren stoßen an ihre Grenzen, da Studierende KI-Tools vergleichsweise einfach zur Täuschung und Leistungserschleichung nutzen könnten. Bezüglich dem Aufspüren und der Einordnung dieser nicht in Eigenleistung erbrachten Leistungen stellen sich ein praktisches und ein definitorisches Problem: Reguläre Plagiatserkennungssoftware erkennt üblicherweise nur Textstellen, die wortwörtlich aus einer anderen Quelle übernommen worden sind, aber keine Texte, die mittels eines Textgenerators erstellt wurden. Auch so genannte ‚AI-Detektoren‘ (wie *GPT-zero* oder *NoGPT*) liefern – schon aus technischen Gründen – bestenfalls Täuschungswahrscheinlichkeiten, keinesfalls aber eindeutige Hinweise oder gar Beweise für tatsächliche Täuschungsversuche. Hinzu kommt, dass Texte von einer KI nach gängigen Plagiatsdefinitionen nicht als zitierfähige Quelle gelten (vgl. Weber-Wulff zit. nach Nehlsen/Fleck 2023, S. 262) und entsprechend ein Text eines KI-Textgenerators „weder als Werk im urheberrechtlichen Sinn noch als zu plagiierte Quelle verstanden“ werden kann und daher daraus auch „keine unmittelbare Zitierpflicht erwächst“ (Nehlsen/Fleck 2023, S. 262 f.). Dennoch handelt es sich bei Texten, die von Textgeneratoren stammen und beispielsweise von Studierenden ohne entsprechende Kennzeichnung in schriftliche (Qualifikations-)Arbeiten übernommen werden, um Täuschungen bezüglich der Autor*innenschaft. Sinnvolle Maßnahmen, um absichtliche oder versehentliche Täuschungsversuche zu vermeiden, sind klare Entscheidungen bezüglich der jeweiligen Lernziele (vgl. u. a. Spannagel 2023a) und hieran angelehnte eindeutige Vorgaben bezüglich erlaubter und nicht erlaubter Hilfsmittel in Prüfungen (vgl. u. a. Nehlsen/Fleck 2023, S. 264; Salden/Lordick/Wiethoff 2023) beziehungsweise Kennzeichnungspflichten in Eigenständigkeitserklärungen schriftlicher Haus-, Seminar- und Qualifikationsarbeiten (vgl. u. a. Abt. Medienpädagogik der PHL 2023; Spannagel 2023b; Weißels/Gottschalk 2023). Nach wie vor erfordern Vorgaben bezüglich erlaubter Hilfsmittel sowie Eigenständigkeitserklärungen gewisses Vertrauen, wenn Lehrende nicht zu Kontrolleur*innen werden wollen oder aufgrund der spezifischen Prüfungsform wie Hausarbeiten oder Onlineklausuren gar nicht in der Lage wären, deren Einhaltung zu kontrollieren. Hier können KI-basierte Textgeneratoren in die lange Liste der möglichen ‚Hilfsmittel‘, wie Taschenrechner, Wikipedia, WWW und Social Media, Eltern oder andere Studierende et cetera, eingeordnet werden, die passend zur jeweiligen Aufgabe und zum Lernziel entweder gestattet oder explizit nicht zugelassen sein können. Unterm Strich erfordern KI-basierte Textgeneratoren – wie zahlreiche Medieninnovationen zuvor auch – eine Neubewertung und Anpassung der Lernziele oder Prüfungsmethoden (vgl. u. a. Salden/Lordick/Wiethoff 2023; Spannagel 2023a; Weißels/Gottschalk 2023), um den sich stetig wandelnden Bedingungen gerecht zu werden.

Ich frage mich, ob in diesen nun deutlich erkennbaren Grenzen tradierter Prüfungsmethoden nicht auch eine Chance steckt: Denn waren diese Prüfungsformen, wie beispielsweise Klausuren, nicht ohnehin stets eine Notlösung und daher schon lange überfällig? Inwiefern war ‚bulimisches Auswendiglernen‘ jemals weiterführend oder eine wesentliche Kompetenz? Im nächsten Absatz möchte ich mich noch etwas eingängiger diesen Fragen widmen.

5. Vier Fragen... und vier grundlegende Strategien

Die bisherigen Ausführungen, die ich in der folgenden Übersicht (vgl. Tabelle 2) noch einmal zusammengestellt habe, zeigten, dass die Nutzung von KI-Tools für Individuen und für unsere Gesellschaft umfängliche Potentiale birgt – vor allem für die Arbeitswelt und die Informationsrecherche beziehungsweise unseren Umgang mit Wissen und sogar für den Bildungsbe- reich. Wie bei jeder Medieninnovation sind mit der Nutzung von KI-Systemen aber auch Fra- gen und Problemlagen verbunden, mit denen sich unsere Gesellschaft befassen muss und aus denen auch Bildungsaufgaben erwachsen, sofern wir Menschen diese Technik einsetzen wol- len – aufgrund der mit ihr verbundenen Potentiale stellt sich diese Frage aber nicht: Wir Men- schen *werden* KI einsetzen.

Tabelle 2: Übersicht der Potentiale und Herausforderungen textgenerierender Dialogsysteme im Bildungskontext

	Potentiale	Herausforderungen
geistige (Kreativ-) Arbeit	effizientere Contentproduktion Inspiration und Brainstorming Entlastung von ‚Akkordarbeit‘	Veränderung von Job-Profilen Re-Re-Re-Produktion inkorrekte Aussagen (‚Halluzinationen‘)
Informa- tions- recherche	Fragen und Nachfragen (Evolution indexbasierter Web-Suchmaschinen) aufbereitete Informationen effizientere/präzisere/tiefere Suche	(weitere) Uniformierung von Wissen Zunahme von Desinformation und Echokammern/Filterblasen verborgene/unreflektierte Stereotype
Schule	‚Sparringspartner*in‘ beim Lernen (wie Hausaufgaben oder Nachhilfe) Umgang mit Heterogenität	der Versuchung widerstehen, der Maschine das Denken zu überlassen (unrefl. Verw.) Kompetenz-/Leistungsüberprüfung
Hoch- schule	Unterstützung des Selbststudiums Förderung heterogener Zielgruppen Chance für ein kreatives Um- und Weiterdenken	Tools nur zur Unterstützung nutzen und die Verantwortung für eigene Leistungen sowie die pers. Weiterentwicklung behalten tradierte Prüfungsformen infragestellen

Betrachtet man die Herausforderungen im Weitwinkel, dann zeigt sich in Anbetracht der zu- vor bereits besprochenen Parallele der technischen Unterstützung körperlicher und geistiger Arbeit des Menschen, dass Technik ein Erweiterer und Verstärker ist: Sie erweitert menschi- che Sinne und Möglichkeiten (vgl. auch McLuhan 1964), sie verstärkt bereits seit langer Zeit die physische Arbeit des Menschen (vgl. Abschnitt 4.1) und seit einigen Jahren auch seine geis- tige Arbeit. Dabei hatte der Mensch seine Werkzeuge aber bisher nicht aus der Hand gegeben (vgl. Abschnitt 6). Heute kann KI jedoch selbständig re-produzieren und wird dadurch zu ei- nem sehr performanten Verstärker geistiger Arbeit. KI verstärkt aber nicht unbedingt das Neue, das Kreative oder das Unbekannte, sondern – das liegt in den (bisherigen) Grenzen der Technologie und der Technik (vgl. Ritter/Martinetz/Schulten 1992; Tickle et al. 1998; Wolf- ram 2023) – das Bekannte und die Mehrheitsmeinung und zementiert damit bestehende (Macht-)Verhältnisse. Damit unterscheidet sich KI nicht von anderen Medieninnovationen, wie dem Buchdruck, dem Radio, dem Fernseher oder dem Computer sowie dem Internet. Das

ist insofern hilfreich, da es sich hierbei um Medieninnovationen handelt, mit denen die Menschheit bereits einen sinnvollen Umgang gefunden hat. Das heißt, es ist möglich, aus den bisherigen Erfahrungen zu lernen und grundlegende Strategien zu übertragen und dadurch *auch* den aktuellen Wandel aktiv und selbstbestimmt zu gestalten (vgl. weiterführend auch Knaus 2017, S. 40–43; Knaus 2018, S. 99–103; Knaus 2020b, S. 60–62).

Hierzu möchte ich im Folgenden vier ganz grundsätzliche Lösungsstrategien vorstellen und eine kritische Auseinandersetzung mit unserem tradierten Bildungsbegriff ergänzen (vgl. Abschnitt 5.4), der zum Nachdenken über Allzuselbstverständliches anregen soll.

5.1 Was sollen wir wissen? Erweiterte subjektive (Medien-)Bildung

Geht es um grundlegende Lösungsstrategien, wie dem Ziel, mit einer Medieninnovation selbstbestimmt umgehen zu können, kommen wir unweigerlich auf mein Lieblingsthema zu sprechen: (Medien-)Bildung. Idealerweise in ihrer um digitalisierungsbezogene Aspekte erweiterten Form, die auch KI als Inhaltsaspekt sowie ihre medien- und gesellschaftskritische Reflexion miteinschließt (vgl. u. a. Knaus 2018, S. 99–103; Knaus 2020a, S. 13; Knaus/Schmidt/Merz 2023). Ansätze, die unter Berücksichtigung des digitalen Wandels inhaltliche Erweiterungen oder interdisziplinäre Verknüpfung anregen, gibt es ja zahlreiche – bereits seit vielen Jahren (vgl. u. a. Brinda et al. 2020; Döbeli Honegger/Kuhnt/Zehnder 2013; Herzig 2001; Knaus 2017; Swertz 2017; Volckart 2022) sowie auch in der aktuellen Ausgabe der *LBzM*, in der auch dieser Beitrag erschienen ist (vgl. u. a. Engel/Kerres 2023; Tulodziecki 2023b; Weberstein/Heidenreich 2023).

Und tatsächlich fordern aktuell nicht wenige aus Politik und Wissenschaft die Förderung spezifischer Kompetenzen im Umgang mit KI: eine so genannte ‚KI-Kompetenz‘²⁷. Wie zuvor konstatiert, wird KI unsere Welt – und vor allem unsere Arbeitswelt (vgl. Abschnitt 4.1) – verändern und nicht unwesentliche Bedeutung für unsere Gesellschaft und Kultur(en) haben. Daher liegt es natürlich nahe, dass Kompetenzen im Umgang mit dieser vergleichsweise neuen Entwicklung gefordert werden. Ich frage mich aber, ob es sich bei der Forderung nach ‚KI-Kompetenzen‘ um den passenden Begriff handelt, um zu beschreiben, dass mit der Zunahme KI-gestützter Anwendungen wie (Text-/Bild-)Generatoren sowie Experten- und Dialogsysteme in unserem Alltag, auch eine Förderung besonderer Fertigkeiten und Fähigkeiten verbunden sein sollte.²⁸ Sehr wahrscheinlich ist eine weitere ‚Bindestrichkompetenz‘ in der Sache wenig hilfreich (vgl. Knaus/Tulodziecki 2023, S. 18–20), aber ist so etwas wie ‚KI-Kompetenz‘ überhaupt möglich und für Menschen erreichbar? Natürlich sollten wir Menschen wissen, womit wir es zu tun haben, wenn wir eine KI nutzen: Wir sollten also wissen, was eine KI ist und idealerweise auch, wie sie funktioniert (vgl. Abschnitt 2). Dieses Wissen dient als solide Entscheidungsgrundlage, um einzuschätzen, welche KI-Tools wir wofür nutzen möchten –

²⁷ Wie beispielsweise das deutsche Bundesministerium für Bildung und Forschung – BMBF ([bmbf.de/bmbf/de/bildung/bildungsforschung/wissenschafts-und-hochschulbildung/ki-in-der-hochschulbildung/ki-kompetenzen.html](https://www.bmbf.de/bmbf/de/bildung/bildungsforschung/wissenschafts-und-hochschulbildung/ki-in-der-hochschulbildung/ki-kompetenzen.html)), das Deutsche Forschungszentrum für Künstliche Intelligenz – DFKI ([dfki.de/web/qualifizierung-vernetzung/netzwerke-initiativen/ki-kompetenzzentren](https://www.dfki.de/web/qualifizierung-vernetzung/netzwerke-initiativen/ki-kompetenzzentren)) oder das Projekt Digitales Deutschland des JFF ([digid.jff.de/magazin/kuenstliche-intelligenz/ki-kompetenzen](https://www.digid.jff.de/magazin/kuenstliche-intelligenz/ki-kompetenzen)).

²⁸ Ausführlicher gehe ich dieser Frage im Kurzvideo *Was ist KI-Kompetenz?* der Videoserie *#Medienpädagogik im Schaukelstuhl* (youtu.be/ACcKBP0OAY8) nach.

eine Entscheidung, die Menschen übrigens bei jedem Medieneinsatz treffen (sollten). Der Kompetenzbegriff deutet aber darauf hin, dass Menschen über bestimmte Fertigkeiten und Fähigkeiten verfügen (sollen), um beispielsweise mit einem bestimmten Medium selbstbestimmt umgehen zu können (ausführlicher zum Begriff der Kompetenz und der Performanz vgl. beispielsweise Knaus/Tulodziecki 2023, S. 13 f.). Wie können wir aber im Kontext künstlicher Intelligenz ‚kompetent‘ – im Sinne von *selbstbestimmt* – sein, wenn wir das Gegenüber selbst gar nicht (umfänglich) einschätzen können? Der Begriff ‚large‘ in *large language model* deutet schließlich darauf hin, dass wir es mit für Menschen nicht überschaubaren Datenmengen (vgl. auch ‚Big Data‘) und insgesamt bei KI-Systemen mit Rechenoperationen zu tun haben, die wir Menschen nicht überblicken und nachvollziehen können – vereinfacht gesagt: Wir können mit einer ‚Black Box‘ nicht selbstbestimmt umgehen. Aus diesem Grund habe ich Zweifel, ob der Begriff der ‚KI-Kompetenz‘ nicht irreführend ist, da unklar ist, ob wir die nötigen Fertigkeiten und Fähigkeiten zum Nachvollziehen der ‚Entscheidungen‘ und ‚Handlungen‘ einer KI überhaupt jemals umfänglich erwerben können. Ist ‚KI-Kompetenz‘ dann nicht ein problematischer Begriff, wenn er suggeriert, dass wir etwas erreichen könnten, was tatsächlich für uns Menschen aber *unerreichbar* ist?

5.2 Was ist die Technik? Mit Explainable AI (XAI) und Tinkering Technik verstehen

Hilfreich wäre es – und damit zur zweiten, ergänzenden Lösungsstrategie –, wenn uns KI-Systeme mit *Meta*-Informationen versorgen würden, das heißt sie würden den Nutzenden verraten und verdeutlichen, auf welcher Datenbasis die Ergebnisse und Informationen beruhen (wie beispielsweise die Trainingsdaten bei LLM) und auf welche Weise die Antworten oder Erkenntnisse generiert wurden – etwa wie ein Taschenrechner, der nicht nur das Ergebnis angibt, sondern auch den Rechenweg aufzeigt. Denn die Hintergründe der Wissensgenerierung sind (bisher) oft nicht klar beziehungsweise ausreichend transparent (vgl. u. a. DFG 2023, S. 1 f.). Wie zuvor (v. a. in Abschnitt 4.2) erläutert, haben Algorithmen und KI-Tools nicht zu unterschätzende Auswirkungen auf unsere Recherchestrategien sowie unseren Umgang mit Wissen. Wenn wir einer KI eine so tiefgreifende Aufgabe zubilligen, dann muss sie – zumindest auf Nachfrage – ihre Wissensfindung offenlegen und transparent machen (können), auf welcher (Daten-)Basis und ‚Entscheidungsgrundlage‘ die Antworten und Erkenntnisse beruhen und wie sie entstanden sind: Die KI legt ihre Datenbasis, die Voreinstellungen und Operationswege offen – etwa so, wie Wissenschaftler*innen in ihren Studien Erhebungs- und Auswertungsmethoden so beschreiben, dass andere Wissenschaftler*innen die Erkenntnisse nachvollziehen können oder bei Wiederholung der Studie die gleichen (oder zumindest vergleichbare) Ergebnisse erhalten würden. Vereinfacht gesagt, würde dadurch KI von einer Black Box zu einer *Glass Box*. In der Informatik wird dieser Ansatz als *Explainable AI* (XAI) oder auch *Algorithmic Explainability* bezeichnet – übrigens weder eine neue Idee, noch ein neuer Forschungszweig (vgl. u. a. Tickle et al. 1998; Heesen et al. 2020).

Bei nicht linear programmierten Systemen generell, aber spätestens im Zuge von Deep Learning mittels künstlicher neuronaler Netze (vgl. weiterführend Ritter/Martinetz/Schulten 1992; Heesen et al. 2020) gerät dieser sehr sinnvolle Ansatz allerdings an seine Grenzen, denn wenn Datenmuster zu komplex sind, um sie zu beschreiben, dann ist es leider auch sehr komplex, diese Muster in eine für den Menschen verständliche und entsprechend nachvollziehbare

Erklärung zu übersetzen. Ein weiterer eher pädagogischer Ansatz, den ich als *Tinkering* schon häufiger vorstellte (vgl. u. a. Knaus/Schmidt 2020; Knaus/Schmidt/Merz 2023) und hier ansetzen könnte, sind *editierbare* Algorithmen: Hierbei können die Voreinstellungen – die Kriterien, Stellgrößen und Schwellwerte –, nach denen eine Maschine ‚entscheidet‘, spielerisch beeinflusst werden. Etwa vergleichbar mit den Voreinstellungen eines Navigationssystems, das in Abhängigkeit der Auswahl ‚Autobahnvorrang‘ oder ‚Autobahnen vermeiden‘ alternative Routen vorschlägt. Interessant wäre auch, wenn Nutzer*innen über die Vorgaben, nach denen Algorithmen die Reihenfolge der Posts in der Timeline einer Social-Media-Plattform (wie *Instagram*, *X (Twitter)*, *Meta (Facebook)* oder *TikTok*) auswählt, selbst bestimmen können: D. h. die Nutzer*innen könnten die Relevanz der Faktoren Aktualität, bisherige Likes, Anzahl der Besuche oder Klicks anderer Nutzer*innen sowie die Intensität der bisherigen Interaktion et cetera für die Auswahl des nächsten Beitrags *selbst* editieren oder aber die algorithmisierte Steuerung auch gänzlich deaktivieren. Kurz: Vorteilhaft wäre, die Voreinstellungen von Algorithmen editieren zu können oder das ‚Lernen‘ von Maschinen gezielt steuern zu können und aus den unterschiedlichen Ergebnissen – etwa wie beim *Reverse Engineering* – verstehen zu können, nach welchen Kriterien die Maschine ‚entscheidet‘ (vgl. auch Knaus/Schmidt/Merz 2023, S. 46 f.).

Ich frage mich, ob wahre ‚KI-Kompetenz‘ nicht das ist, was Wissenschaftler*innen machen: Nämlich (theoretische) Hintergründe zu (er-)kennen und Sachverhalte beziehungsweise Thesen auf ihren Wahrheitsgehalt prüfen. Aber dazu später mehr. Zunächst noch kurz zum dritten, ergänzenden Lösungsansatz.

5.3 Was darf ich hoffen, wenn subjektive Medienbildung und XAI an ihre Grenzen stoßen?

Die zuvor beispielhaft vorgestellten Ansätze erfordern eine gewisse Offenheit und Transparenz der Systeme; auch der Ansatz Explainable AI (XAI) macht beabsichtigte Entscheidungen und entsprechende Voreinstellungen der Tool- und Plattformanbieter*innen transparent und nachvollziehbar. Diese nötige prinzipielle Offenheit und Transparenz scheitert mitunter an *Möglichkeiten* (wie zuvor am Beispiel von XAI dargelegt), aber auch nicht selten am *Willen* der Anbieter*innen, die in Algorithmen beziehungsweise im Code eingeschriebenen Strategien offenzulegen. Strategien, die den kommerziellen Anbieter*innen – die als Unternehmen verständlicherweise Gewinne erwirtschaften wollen – Vorteile am (Werbe-)Markt einbringen. Hier stellt sich mir die Frage, warum für uns (inzwischen) sehr relevante und selbstverständliche Dienste, wie Editoren, Messenger und Social-Media-Plattformen, eigentlich nichts kosten dürfen. Diese Erwartung stammt aus einer Zeit, in der alle Dienste im Netz – aufgrund seiner nicht-kommerziellen Historie – für alle Nutzenden frei verfügbar sein sollten. Dieses an sich erfreuliche Erbe führte jedoch im Zuge der Teil-Kommerzialisierung des Netzes und seiner Dienste dazu, dass Nutzende mit Werbung und immer neuen Strategien der hierfür nötigen algorithmisierten ‚Aufmerksamkeitssteuerung‘ konfrontiert sind – inzwischen sogar mit De-

signentscheidungen, die von den Anbieter*innen absichtlich so gestaltet sind, dass sie die Nutzer*innen zu etwas verleiten, was sie nicht beabsichtigten.²⁹ Werbung und gezielte (von Algorithmen oder einer KI gesteuerte) Aufmerksamkeitsteuerung³⁰ mag nervig sein, aber diese so genannten *Dark Patterns* berühren zudem auch eine rechtliche Grauzone, mit der sich auch die Europäische Kommission unter anderem im Rahmen ihrer Verbraucheragenda befassen möchte (vgl. EU 2020; EDPB 2022).

Einer der ersten politischen Impulse auf mit Unkenntnis einhergehender Überforderung ist nicht selten das Verbot. Auch nach der öffentlichkeitswirksamen Vorstellung von ChatGPT war dieser Impuls zu beobachten: So wurde beispielsweise in Italien die Nutzung von ChatGPT verboten, der Zugang zum Dienst aus wenig nachvollziehbaren Gründen blockiert (vgl. u. a. Schubert 2023). Das Verbot wurde inzwischen wieder aufgehoben. Als weitere Beispiele kann ich hier noch einmal das bereits erwähnte ‚KI-Moratorium‘ (vgl. Future of Life Institute 2023) nennen, das einige Wissenschaftler*innen und Eigner*innen von Tech-Konzernen forderten (vgl. weiterführend u. a. Anderl 2023; Krempf 2023; Knaus/Tulodziecki 2023, S. 7 und 12). Interessant ist, dass es sich bei diesen Reaktionen um sehr übliche und regelmäßig wiederkehrende Impulse handelt, die bei keiner Medien- oder Technikinnovation der Geschichte ausblieben (vgl. Kümmel/Scholz/Schumacher 2004, S. 8): Dazu gehören unter anderem der Dialog zwischen Thamus und Theuth zur Schrift aus Platons *Phaidros* (vgl. Heitsch 1993, S. 60–65), Warnungen vor dem ‚Delirium furiosum‘ bei der Nutzung der (Eisen-)Bahn³¹, der

²⁹ Diese Designentscheidungen von Benutzerschnittstellen (wie Webseiten oder auch in Apps), die absichtlich so gestaltet sind, dass sie die Nutzer*innen zu etwas verleiten sollen, was sie nicht beabsichtigten, werden als *Digital Nudging* beziehungsweise *Dark Pattern* bezeichnet. Solche Dark Patterns ‚funktionieren‘, da Menschen eine Entscheidung (wie den Kauf eines bestimmten Produkts oder einer Dienstleistung) grundsätzlich nicht nur von rationalen Kriterien abhängig machen, sondern diese stets auch von emotionalen Aspekten mitbestimmt werden. Typische Beispiele für ‚Nudging‘ mittels Dark Patterns finden sich tagtäglich im Netz: So findet man beispielsweise in zahlreichen Kundenportalen vergleichsweise ganz einfach Möglichkeiten zum Upgrade, aber keinen einfachen Weg zum Downgrade oder zur Kündigung eines Abos oder eines Dienstes. Oder niedrige Preise, denen versteckte Kosten folgen, wie bei einer Flugbuchung (Hidden Costs). Oder auch künstlich erzeugter Stress beim Kauf oder einer Buchung durch den Hinweis, dass dieses Angebot in wenigen Stunden oder Minuten abläuft und dann nicht mehr verfügbar ist. Auch Trickfragen, wie das ‚Cookie Consent Tricking‘, bei denen Nutzende eine Checkbox *aktivieren* müssen, um *keine* Werbung oder einen Newsletter zu erhalten: „Klicke hier um keine Werbung zu erhalten.“ Es gibt zahlreiche weitere Beispiele (vgl. u. a. Brignull 2022) – auch in Social-Media-Plattformen (vgl. Seemann 2021), die ein selbstbestimmtes Medienhandeln erschweren (vgl. weiterführend auch Kammerl et al. 2023, S. 44–54). Etwas ausführlicher befasse ich mich mit Dark Patterns im Kurzvideo *Was sind Dark Patterns (Und warum sie zeigen, dass der digitale Wandel eine gesellschaftliche Normierungs- und Gestaltungsaufgabe ist...)* der Videoserie *#Medienpädagogik im Schaukelstuhl*: youtu.be/dMePixWg2oA.

³⁰ Eine vergleichsweise einfache Lösung zur Eindämmung dieser Trickereien, (Micro-)Targeting und überbordender Werbung wäre, wenn Nutzer*innen und Unternehmen bereit wären, für die ihnen wichtigen Dienste zu bezahlen. Auch hier zeigt die Entwicklung vom werbefinanzierten Privatfernsehen hin zum kostenpflichtigen Streamingdienst retrospektiv aufschlussreiche Parallelen auf (vgl. Abschnitt 4). Diese Parallele zeigt aber leider auch, dass selbst professionalisierte Geschäftsmodelle die Netzökonomie bedauerlicherweise nicht zuverlässig vom ‚Digital Nudging‘ abhalten (vgl. Chaudhary et al. 2022). Ein denkbarer Ausweg wäre, wenn größere Verbände wie Unternehmen, Behörden oder Länder eigene (oder so genannte mandantenfähige) Systeme und Dienste nach *eigenen Vorgaben* entwickeln (lassen) und betreiben.

³¹ Im 18. Jahrhundert warnten Experten davor, dass Geschwindigkeiten über 30 Stundenkilometer in der damals neuen Eisenbahn gesundheitsschädlich sein könnten. Die Menschen misstrauten der Bahn, da sie schneller als alle bisherigen Verkehrsmittel war – zumal das Pferd zu dieser Zeit als Maßstab galt. Als George Stephenson 1825 die erste Eisenbahnstrecke zwischen Manchester und Liverpool beantragte, wurde aufgrund dieser Bedenken vom britischen Unterhaus ein Gutachten der Pariser Akademie der Wissenschaften eingeholt. Dieses Gutachten warnte vor der schnellen Bewegung, da sie

Lesesucht³² oder den schädlichen Auswirkungen des Kinos für die Phantasie (vgl. u. a. Döblin 1909, S. 191 f.; Pfemfert 1992 oder weiterführend auch Kümmel/Scholz/Schumacher 2004 oder Wessler 1997). Medien- und Technikinnovationen provozieren also wiederkehrend Befürchtungen und entsprechende Warnungen oder Verbote. Dabei wissen wir doch eigentlich aus der empirischen Rechtssoziologie, der Sozialpsychologie und auch der Kulturgeschichte, dass übereilte Verbote selten erfolgreich waren (vgl. u. a. Welskopp 2010) und vielmehr Aufklärung über Möglichkeiten und Gefahren sowie das Erlernen eines sinnvollen und selbstbestimmten Umgangs wesentlich ist – ob es nun um Medien wie Streamingdienste und Social Media, Genussmittel wie Schokolade oder den Konsum von Rauschmitteln geht. In Bezug auf die Mediennutzung kommt die in der Medienpädagogik traditionell sehr wesentliche *Selbstbestimmung* in der Nutzung von Medien, Diensten und Plattformen aus den zuvor benannten Gründen an ihre Grenzen (vgl. Knaus 2018, S. 98 f.; Knaus 2020a, S. 11; Kammerl et al. 2023; S. 51–54) – eine Analyse, die auch die anderen Beiträge in der Ausgabe der vorliegenden *LBzM* stützen (vgl. Barberi/Iske 2023; Dander 2023; Engel/Kerres 2023; Tulodziecki 2023b; Weberstein/Heidenreich 2023 sowie Knaus/Tulodziecki 2023, S. 18 f. jeweils in dieser Ausgabe). Denn ein Individuum kann sich noch so überlegt im Netz beziehungsweise im WWW oder in Sozialen Netzwerken bewegen, wenn Webseiten, unscheinbare Apps³³, ‚intelligente Lautsprecher‘³⁴ oder Kameras im Öffentlichen Raum persönliche Daten sammeln und weitergeben, ist es mit der ‚Datensouveränität‘ passé (vgl. Knaus 2018, S. 98 f.). Seit dem Skandal um *Facebook* (heute: *Meta*) und *Cambridge Analytica* (vgl. weiterführend u. a. Confessore 2018; Hansen 2018), der in der Öffentlichkeit zu einer etwas sensibleren Wahrnehmung von Datenschutzvergehen führte, erfahren wir wöchentlich von weiteren Beispielen der versteckten oder unerlaubten Nutzung von Daten aus Apps oder Social-Media-Plattformen, beispielsweise zum Zwecke von (Micro-)Targeting.

Aus diesem Grund sind – neben einer um digitalisierungsbezogene Aspekte erweiterten (subjektiven) Medienbildung (vgl. auch Abschnitt 3) und XAI – weitere Normierungsbestrebungen, Ordnungsversuche (vgl. u. a. EU 2019; EU 2020), die Diskussion und Klärung ethischer Fragen wie auch ethisch-normativer Leitlinien (vgl. u. a. Dabrock 2023; Deutscher Bundestag

möglicherweise eine Gehirnerkrankung namens ‚Delirium furiosum‘ verursachen könnte. Wenige Jahre später (1838) soll das Bayerische Obermedizinalkollegium zu einem ähnlichen Ergebnis gekommen sein: „Die schnelle Bewegung muß bei den Reisenden unfehlbar eine Gehirnkrankheit, eine besondere Art des delirium furiosum erzeugen. Wollen aber dennoch Reisende dieser gräßlichen Gefahr trotzen, so muß der Staat wenigstens die Zuschauer schützen, denn sonst verfallen diese beim Anblick des schnell dahinfahrenden Dampfwagens genau derselben Gehirnkrankheit“ (Rössler 1983, S. 11).

³² Es ist kaum zwei Jahrhunderte her, dass das Lesen Auslöser und Gegenstand äußerst kritischer Diskussionen war, deren Inhalte bis hin zu einer diagnostizierten ‚Lesesucht‘ und entsprechenden Leseverboten reichten. Ein Beispiel ist Goethes Werk *Die Leiden des jungen Werther* – heute ein Klassiker der Weltliteratur. Damals stand das Buch allerdings im Verdacht, Menschen nicht nur „von der Arbeit abzulenken“, sondern ihnen auch dauerhaft zu schaden (vgl. u. a. Bollmann 2005, S. 25; Hoche 1794 und weiterführend auch Wessler 1997).

³³ Zur Nutzung des Messengers *WhatsApp* ist das Freigeben des persönlichen Adressbuchs erforderlich. Durch diese Freigabe werden auch Namen und Kontaktdaten von Personen übertragen, die sich möglicherweise gegen die Nutzung des Dienstes entschieden haben: Personen, die im persönlichen Adressbuch eines anderen WhatsApp-Kontakts gespeichert sind, können nicht souverän und fallbezogen über die Speicherung und Weitergabe ihrer persönlichen Daten entscheiden (vgl. weiterführend Knaus 2018, S. 98 f.; Knaus 2020a, S. 11 f.).

³⁴ So genannte ‚Smart Speaker‘, wie *Amazon-Echo*, *Apple HomePod* oder *Google-Home*.

2020; Deutscher Ethikrat 2023; EDPB 2022) nötig. Erforderlich ist überdies eine multidisziplinäre Forschung zur gesellschaftlich-kulturellen Bedeutung von KI, da diese (politischen) Entscheidungen neben einer persönlichen Expertise stets auch einen umfassenden kritischen Blick auf die gesellschaftlichen und individuellen Folgen erfordern. Vereinfacht gesagt: Wir müssen wissen, womit wir es zu tun haben und was dies mit uns und unserer Gesellschaft macht.

5.4 Was sollen wir tun? Ein kritischer Blick auf unseren Bildungsbegriff

KI wird im Bildungskontext – vor allem mit Blick auf die Schule und in der Hochschule – aktuell primär als Bedrohung diskutiert: Textgenerierende KI würde die Kulturtechnik des Schreibens bedrohen (vgl. u. a. Dabrock 2023). Und tatsächlich ist das Schreiben nicht nur ein Aufzeichnen von Schriftzeichen oder eine Kommunikationspraxis, sondern auch ein nicht zu unterschätzendes Denkwerkzeug (vgl. u. a. Bereiter 1980). Wurde aber nicht just diese Kulturtechnik schon mehrfach durch Medieninnovationen angefragt (vgl. Faßler/Halbach 1998; Heitsch 1993, S. 60–65; Scholz 2004, S. 19–27)? Tatsächlich kommunizieren wir heute weiterhin intensiv in Schriftform – zum Beispiel mittels (Instant-)Messengern wie *WhatsApp* (vgl. u. a. MPFS 2021a, S. 38; MPFS 2021b, S. 26; MPFS 2022, S. 26 f.). Auch die tradierten Prüfungsformen in Schule und Hochschule sind für nicht wenige Lehrende in Schulen und Hochschulen ein Grund zur Sorge: KI-basierte Dialogsysteme und Textgeneratoren können Antworten für Klausuren liefern, die weder von Prüfer*innen noch von üblicher Plagiaterkennungssoftware identifiziert werden können (vgl. Abschnitt 4.4). Um diesen vergleichsweise neuen Herausforderung zu begegnen, wurden bereits spezielle Tools (wie *GPTzero* oder *No-GPT*) entwickelt, die Text(teil)e, die mittels einer textgenerierenden KI erstellt wurden, aufspüren und erkennen können – doch auch diese ‚AI-Detektoren‘ liefern bestenfalls Täuschungswahrscheinlichkeiten und keine Beweise für Täuschungsversuche. Steckt aber nicht just in dieser Herausforderung auch eine Chance für den Bildungsbereich, nicht nur das Reproduzierbare zu prüfen, sondern Aufgaben (auch für Qualifikationsarbeiten) so zu stellen, das Schüler*innen und Studierende Wissen nicht re-produzieren müssen – etwas, was Maschinen ohnehin allein strukturell einfach besser können. Sollten Lernende nicht vielmehr dazu angeregt werden, (eigenes) ‚Wissen zu schaffen‘, beispielsweise auf Grundlage eigener Forschungsfragen und mittels empirischer Forschung? Haben nicht zahlreiche Lernforscher*innen – gerade die, die konstruktivistische Lerntheorien vertreten – schon lange auf die sinnvolle Kombination von Instruktion (‚learning when the answer is known‘) und Konstruktion (‚learning when the answer is not known‘) hingewiesen (vgl. u. a. Siebert 2005, S. 33; Reich 2008)? Wenn wir uns aber anschauen, wie in Schule und Hochschule üblicherweise gelernt wird, dann ist es häufig so, dass Lernende (nach wie vor) primär mit bekannten Antworten konfrontiert werden. Dabei wäre es doch viel interessanter, wenn wir – gerade in den höheren Klassenstufen und der Hochschule – mehr Dinge lernen würden, deren Antwort wir noch *nicht* kennen – ich meine damit das entdeckende, forschende Lernen (vgl. auch BAK 1970). Die Instruktion – also die Aufgabe, Schüler*innen an das (Welt-)Wissen heranzuführen – bleibt natürlich weiterhin relevant, da sie die Basis des Weiterdenkens und darauf aufbauender (Forschungs-)Fragen ist, und das kreative Zusammenspiel von Instruktion und Konstruktion bleibt eine wichtige Wechselbeziehung in Lernvorgängen. Trotz der Relevanz der Konstruk-

tion wird auf vorgefertigte Antworten nicht selten auch aus lehr- und lernökonomischen Beweggründen zurückgegriffen, da nur diese auch effizient geprüft werden können: Klausuren eignen sich zum effizienten ‚Abprüfen‘ großer Klassenverbände und in Massenstudiengängen in besonderer Weise. Der Verdacht drängt sich auf, dass Klausuren primär aus Effizienzgründen eingeführt wurden. Dieser Verdacht erhärtet sich durch die naheliegende Vermutung, dass sich wohl kaum Schüler*innen oder Studierende an außergewöhnlich kluge Klausurfragen erinnern werden, die sie zu besonderen Erkenntnissen inspiriert haben. Nur wenige Lehrende führten wohl eine Klausur als Prüfungsleistung ein oder hielten aus Überzeugung an der Lernförderlichkeit dieser Prüfungsform daran fest. Unglücklicherweise verleitet diese Prüfungsform überdies auch dazu, im Unterricht sowie in Seminar und Vorlesung re-produzierbares Wissen anzubieten beziehungsweise es verleitet die Lernenden dazu, genau diese gesicherten Standardantworten gezielt zu identifizieren („Ist das prüfungsrelevant?“) und (auswendig) zu lernen. Die Prüfungsform beeinflusst also nicht nur die Prüfungsvorbereitung, sondern auch den Unterricht beziehungsweise die Lehrveranstaltung selbst, dabei sind Klausuren doch nur Notlösungen, da mündliche Prüfungen für alle Beteiligten deutlich zeitaufwändiger wären (vgl. Marchand 2023). An diesem Kapazitätsproblem hat sich bis heute nichts geändert, und doch werden einige Stimmen lauter, die behaupten, dass die einzig täuschungssichere Prüfung in ‚KI-Zeiten‘ die mündliche Prüfung sei (vgl. Salden/Lordick/Wiethoff 2023, S. 16 f.; Spannagel 2023a; kritisch dazu u. a. Marchand 2023). Wenn wir davon ausgehen, dass für Prüfungen keine zusätzlichen (personellen) Ressourcen geschaffen werden, dann würde künftig mehr Lernzeit zum Abprüfen re-produzierbaren Wissens aufgewendet. Ich frage mich, ob nicht genau das Gegenteil sinnvoller wäre. Und ich frage mich weiter, ob es nicht kreativitäts- und lernförderlich sein könnte, wenn wir die KI-Dialogsysteme, die uns im Alltag Antworten geben (vgl. Wolfram 2023), in pädagogischen Kontexten als ‚Fragemaschinen‘ nutzen, die uns Lehrende dabei unterstützen, Studierenden und Schüler*innen kluge und anregende Fragen zu stellen? Und mit diesen Fragen kreativ (ver-)stören (vgl. Knaus 2013) – wie es meine Co-Autor*innen in der Genese dieses Texts auch mit mir machten (vgl. Abschnitt 1).

Die sich wiederholende Aufgabe im Umgang mit den Herausforderungen der jeweils neuen Technik- und Medieninnovationen bestand stets darin, einen angemessenen Umgang mit den medialen und technischen Innovationen zu erlernen. Dies war geprägt von sowohl individuellen als auch gesellschaftlichen Entscheidungen, wie Bildungsentscheidungen, gezielter (pädagogischer) Förderung, aber nicht zuletzt auch durch das Festlegen von Regeln und Normen (vgl. Abschnitt 5.3). So haben Computer und Internet die Arbeitswelt nachhaltig geprägt; und auch Rechenmaschinen, wie der Taschenrechner, erleichtern ganz selbstverständlich Kalkulationsaufgaben im Alltag. Um aber Heranwachsende in die Lage zu versetzen, die Ergebnisse eines Rechners auf Plausibilität zu prüfen, lernen sie im Mathematikunterricht nicht nur dessen Bedienung (mehr zum Begriff der ‚Be-dienung‘ einer Maschine im nächsten Abschnitt), sondern vielmehr mathematische Sachverhalte, Rechenregeln, Rechenstrategien und die jeweiligen Rechenwege zu *verstehen*. In diesem einfachen Beispiel stecken einige Erfahrungen aus bisherigen gesellschaftlichen Transformationsprozessen und entsprechend auch viel Anknüpfbares, wie wir auch in Zukunft mit den aktuellen Entwicklungen strukturell umgehen können.

6. Kolleg*in KI? – (m)ein Antwortversuch

Abschließend möchte ich einen Antwortversuch auf die im Rahmen der Ringvorlesung aufgeworfene Frage (vgl. Abbildung 2) unternehmen: Soll KI unsere Kolleg*in sein?



Abbildung 2: Ausschnitt des Plakats der Ringvorlesung *Transformation@Work: Kollaboration mit Kolleg*in KI* des Fraunhofer-Instituts für Arbeitswirtschaft und Organisation (IAO) und dem Institut für Arbeitswissenschaft und Technologiemanagement (IAT) der Universität Stuttgart

Kolleg*innen eines Arbeits- oder Projektteams verfolgen in der Regel ein gemeinsames Ziel und erfüllen zum Erreichen des Ziels bestimmte Aufgaben, die sie dem gesetzten Ziel näherbringen. Die Frage, ob eine Arbeitsteilung zwischen Mensch und Maschine sinnvoll ist, führt daher zur Frage, welche Beiträge die beiden Kolleg*innen – der Mensch und die Maschine – leisten könnten: Welche Aufgaben können künftig Maschinen übernehmen, mit dem Ziel, uns zu unterstützen? Und welche Aufgaben können oder sollen in menschlicher Hand und Verantwortung bleiben? Wie die vorherigen Ausführungen und Beispiele zeigten, sind Maschinen mit Blick auf nicht-physische Arbeit besonders gut im Erledigen wiederkehrender Aufgaben, der Re-Produktion, dem Regelbefolgen, dem Speichern und Wieder-Finden; während der Mensch zu innovativen Gedanken, dem Erfinden und zum kreativen Out-of-the-Box-Denken in der Lage ist (vgl. weiterführend Knaus/Merz/Junge 2023) und seine Verantwortung für Care-Aufgaben, Erziehung und Bildung nicht aus der Hand geben sollte. Eine produktive Zusammenarbeit im Sinne einer Reduzierung wiederkehrender (Routine-)Aufgaben mit dem Ziel Freiräume zu schaffen, erscheint also möglich und sinnvoll – zumal der Blick in die jüngere Geschichte zeigt, dass die *physische* Arbeitsteilung von Mensch und Maschine genau diese Erwartungen eingelöst hat (vgl. Abschnitt 4.1).

Eine nicht unwesentliche Frage ist aber: Wer der Kolleg*innen unterstützt eigentlich wen? Anders gesagt: Wer hat in der Kollaboration eigentlich ‚das Sagen‘, wer bestimmt die Regeln der Zusammenarbeit und wer steht im Mittelpunkt? Interessanterweise war und ist es oft die Technik! Ja, tatsächlich: Die Technik steht in kollaborativen Arbeitskontexten nicht selten im Mittelpunkt (vgl. u. a. Shneiderman 2020). Diese Merkwürdigkeit ist historisch gewachsen

und im Wesentlichen ein Erbe aus der Zeit der Industrialisierung. Sehr aufschlussreich bezüglich der bisherigen Arbeitsteilung von Mensch und Maschine und der Aufdeckung dieser historisch gewachsenen Merkwürdigkeit kann die kritische Auseinandersetzung mit dem deutschen Begriff des ‚Bedienens‘ sein: Was bedeutet es eigentlich wortwörtlich, wenn wir sagen, dass Menschen beispielsweise ein digitales Medium ‚be-dienen³⁵‘ können sollen? Dieser Begriff scheint aus einer anderen Zeit zu stammen und hat daher recht wenig mit dem zu tun, wie wir heute Medien und Technik nutzen – zumindest wie wir sie nutzen *sollten*, nämlich selbstbestimmt und zur Erreichung persönlicher Ziele. Bedienen deutet ja darauf hin, dass der Mensch, der die Maschine *be-dient*, sich um die Maschine kümmert und bemüht. In diesem Fall steht die Maschine im Mittelpunkt. In der Zeit der Industrialisierung war das nötig, damit ‚das Band‘ – als Metapher für die Produktion – keinesfalls stillsteht. Die Maschine musste laufen – sie stand im Zentrum –, da sie seinerzeit deutlich teurer war als menschliche Arbeit. Heute haben sich die Bedingungen grundlegend verändert, aber diese historisch gewachsenen Denkweisen scheinen bis heute fortzubestehen, wenn technische Vorgaben menschliches Handeln determinieren (vgl. weiterführend Knaus 2020b, S. 19–22). Sollten Maschinen heute nicht eher uns Menschen (be-)dienen? Immerhin sind wir Menschen es, die Maschinen gestalten können, die uns unterstützen (Knaus 2020b, S. 47 f. und 61). Der Mensch stünde dann (wieder) im Mittelpunkt (vgl. Shneiderman 2020), wenn Technik – und entsprechend auch KI-Systeme – den Menschen bei der Erreichung seiner Ziele unterstützen.

Am Ende eines aufwändigen mensch-maschineübergreifenden Diskussions- und Entwicklungsprozesses (vgl. Abschnitt 1), der diesen ausführlichen Beitrag als Ergebnis zustande brachte, passt mein persönliches Fazit auf die im Rahmen dieser Ringvorlesung aufgeworfene Frage nun in einen kurzen Satz: Als Werkzeug ja, als Partner*in nein und als Kollege*in nur dann, wenn der Mensch selbstbestimmt gestalten kann und im Mittelpunkt steht.

Credits

Herzlichen Dank an Frau Kollegin Professorin Katharina Hölzle (Fraunhofer IAO; Institutsleiterin des IAT der Universität Stuttgart) und ihre Mitarbeiterin Lisa Kurz für die freundliche Einladung zum Vortrag an der Universität Stuttgart und die damit verbundenen Inspirationen zu diesem Beitrag. Umfänglicher Dank gilt nicht nur der künstlichen Intelligenz, die als KI-basierte Sprachmodelle einige Zeilen zu diesem Text beisteuerten, mich als Dialogsysteme unterstützten und dabei nicht wenige Ressentiments bestätigten, aber mich auch immer wieder positiv überraschten, sondern auch der kreativ-kritischen Humanintelligenz der Redaktion der *LBzM*, den Kolleg*innen Olga Merz und Thorsten Junge. Und last but not least danke ich sehr herzlich meiner Doktorandin Jennifer Schmidt und *Midjourney* für kreative Prompts, interessanten Austausch zu den Möglichkeiten bildgenerierender KI und beeindruckende Bilder für meine Vortragsfolien.

³⁵ Weiterführende Hinweise lassen sich im kurzen Video „Warum genügt es nicht, dass wir Menschen Medien nur ‚bedienen‘ können?“ der Videoserie *#MedienpaedImSchaukelstuhl*: youtu.be/SMFtkE6ShSc finden.

Literatur

- Abt. Medienpädagogik der PHL (2023): Hinweise und Stylesheet für das Verfassen wissenschaftlicher Qualifikationsarbeiten in der Medienpädagogik. ph-ludwigsburg.de/fakultaet-1/institut-fuer-erziehungswissenschaft/medienpaedagogik/studium-und-lehre/qualifikationsarbeiten.
- Anderl, Sibylle (2023): Brauchen wir eine KI-Pause? faz.net/aktuell/wissen/computer-mathematik/unternehmer-und-forscher-fordern-entwicklungspause-fuer-ki-18793216.html.
- Attwell, Graham (2007): The Personal Learning Environments – The future of eLearning? In: *eLearning Papers*, 2 (1), pp. 1–8. [researchgate.net/publication/228350341_Personal_Learning_Environments-the_future_of_eLearning](https://www.researchgate.net/publication/228350341_Personal_Learning_Environments-the_future_of_eLearning).
- BAK – Bundesassistentenkonferenz (1970): *Forschendes Lernen – Wissenschaftliches Prüfen*, Bonn: BAK.
- Baggaley, Jon (2013): MOOC rampant. In: *Distance Education*, 34 (3), p. 368–378, doi.org/10.1080/01587919.2013.835768.
- Bajohr, Hannes (2022): *Schreibenlassen – Texte zur Literatur im Digitalen*. Berlin: August.
- Barberi, Alessandro/Iske, Stefan (2023): Zur ‚Spontaneität des Menschen‘ – Sprachphilosophische und anthropologische Grundlagen der Medienpädagogik in Dieter Baackes ‚Kommunikation und Kompetenz‘. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 23/2023. doi.org/10.21240/lbzm/23/03.
- Bass, Dina (2023): Microsoft Invests \$10 Billion in ChatGPT Maker OpenAI. [bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai](https://www.bloomberg.com/news/articles/2023-01-23/microsoft-makes-multibillion-dollar-investment-in-openai).
- Beranek, Angelika (2020): Beyond the Black Box – Was steckt hinter dem Interface? Programmierte Werte und die Rolle der Medienpädagogik. In: Knaus, Thomas/Merz, Olga (Hrsg.): *Schnittstellen und Interfaces – Digitaler Wandel in Bildungseinrichtungen* (Bd. 7). München: kopaed, S. 73–92.
- Bereiter, Carl (1980): *Development in Writing*. In: Gregg, Lee/Steinberg, Erwin (Hrsg.): *Cognitive processes in writing*. London: Routledge, pp. 73–93.
- Bollmann, Stefan (2005): *Frauen, die lesen, sind gefährlich*. München: Sandmann.
- Brandt, Mathias (2023): Wie lange brauchen Online-Dienste, um eine Million Menschen zu erreichen. de.statista.com/infografik/29195/zeitraum-den-online-dienste-gebraucht-haben-um-eine-million-nutzer-zu-erreichen.
- Brignull, Harry (2022): *What is deceptive Design?* deceptive.design.
- Brinda, Torsten/Brüggen, Niels/Diethelm, Ira/Knaus, Thomas/Kommer, Sven/Kopf, Christine/Mis-somelius, Petra/Leschke, Rainer/Tilemann, Friederike/Weich, Andreas (2020): Frankfurt-Dreieck zur Bildung in der digital vernetzten Welt – Ein interdisziplinäres Modell. In: Knaus, Thomas/Merz, Olga (Hrsg.): *Schnittstellen und Interfaces – Digitaler Wandel in Bildungseinrichtungen* (Bd. 7 der fraMediale-Reihe). München: kopaed, S. 157–167. doi.org/10.25656/01:22117.
- Bruce, David L. (2008): Visualizing literacy – Building bridges with media. In: *Reading & Writing Quarterly*, 24, pp. 264–282. doi.org/10.1080/10573560802004126.
- Chaudhary, Akash/Saroha, Jairvat/Monteiro, Kyzyl/Forbes, Angus G./Parnami, Aman (2022): ‘Are You Still Watching?’ Exploring Unintended User Behaviors and Dark Patterns on Video Streaming Platform. *DIS 2022*, pp 776–791. doi.org/10.1145/3532106.3533562.
- Confessore, Nicholas (2018): Cambridge Analytica and Facebook – The scandal and the fallout so far. [nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html](https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html).
- Conroy, Gemma (2023): Scientists used ChatGPT to generate an entire paper from scratch — but is it any good? [nature.com/articles/d41586-023-02218-z](https://www.nature.com/articles/d41586-023-02218-z).
- Coy, Wolfgang (1995): *Automat – Werkzeug – Medium*. In: *Informatik Spektrum* 18 (1), S. 31–38.
- Cuban, Larry (1986): *Teachers and machines – The classroom use of technology since 1920*. New York: Columbia University.

- Dabrock, Peter (2023): So lässt sich ChatGPT verantworten. spiegel.de/netzwelt/chatgpt-so-laesst-sich-kuenstliche-intelligenz-verantworten-gastbeitrag-a-d89746ff-a263-4a70-a6d2-7029bb45b7ac.
- Dander, Valentin (2023): Politische Medienbildung und Netzpolitik. Eine Re-Aktualisierung des ‚Diskurses der Informationsgesellschaft‘. In: Ludwigsburger Beiträge zur Medienpädagogik – LBzM, 23/2023. doi.org/10.21240/lbzm/23/06.
- Derrida, Jacques (2001): Die unbedingte Universität. Frankfurt am Main: Suhrkamp.
- Deutscher Bundestag (2020): Bericht der Enquete-Kommission Künstliche Intelligenz – Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale. Bundestagsdrucksache 19/23700. Berlin. dserver.bundestag.de/btd/19/237/1923700.pdf.
- Deutscher Ethikrat (2023): Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Vorabfassung der Stellungnahme. Berlin: Geschäftsstelle der Deutschen Ethikrats. ethikrat.org/fileadmin/Publikationen/Stellungnahmen/deutsch/stellungnahme-mensch-und-maschine.pdf.
- DFG (2023): Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG. dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2023/230921_stellungnahme_praesidium_ki_ai.pdf.
- Döbeli Honegger, Beat (2023): ChatGPT – der iPhone-Moment des maschinellen Lernens? In: info 7, 23 (1). beat.doebe.li/publications/2023-doebeli-honegger-chatgpt-info7.pdf oder info7.de/info7_2023-1.pdf.
- Döbeli Honegger, Beat/Kuhnt, Beate/Zehnder, Carl August (2013): Informatik, ICT und Medienbildung, In: Kohlas, Jürg/Schmid, Jürg/Zehnder, Carl August: Informatik@gymnasium. Zürich: NZZ.
- Döblin, Alfred (1909): Das Theater der kleinen Leute. In: Das Theater, 1, S. 191–192.
- EDPB – European Data Protection Board (2022): Guidelines 3/2022 on Dark patterns in social media platform interfaces – How to recognize and avoid them. edpb.europa.eu/system/files/2022-03/edpb_03-2022_guidelines_on_dark_patterns_in_social_media_platform_interfaces_en.pdf.
- Eloundou, Tyna/Manning, Sam/Mishkin, Pamela/Rock, Daniel (2023): GPTs are GPTs – An Early Look at the Labor Market Impact Potential of Large Language Models. doi.org/10.48550/arXiv.2303.10130.
- Engel, Juliane/Kerres, Michael (2023): Bildung in der Nächsten Gesellschaft – Eine postdigitale Sicht auf neue Formen der Subjektivierung. In: Ludwigsburger Beiträge zur Medienpädagogik – LBzM, 23/2023. doi.org/10.21240/lbzm/23/04.
- Ernst, Anna (2023): KI mit Soße. sueddeutsche.de/medien/burda-lisa-kochen-backen-rezepte-ki-leser-1.5855586.
- EU – European Commission (2019): Ethics guidelines for trustworthy AI. digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.
- EU – European Commission (2020): New Consumer Agenda. ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_20_2069/IP_20_2069_EN.pdf.
- Faßler, Manfred/Halbach, Wulf (1998): Geschichte der Medien. Stuttgart: UTB.
- Felten, Michael (2010): Auf die Lehrer kommt es an! Für eine Rückkehr der Pädagogik in die Schule, Gütersloh: Gütersloher Verlagshaus.
- Fend, Helmut (1979): Gesellschaftliche Bedingungen schulischer Sozialisation, Weinheim: Beltz.
- Ferster, Bill (2014): Teaching machines – Learning from the intersection of education and technology. Baltimore: Johns Hopkins University Press.
- Freudenstein, Reinhold (1975): Unterrichtsmittel Sprachlabor – Technik, Methodik, Didaktik. Bochum: Kamp.

- Future of Life Institute (2023): Pause Giant AI Experiments – An Open Letter. futureoflife.org/wp-content/uploads/2023/05/FLI_Pause-Giant-AI-Experiments_An-Open-Letter.pdf.
- Generative Pretrained Transformer, Gpt/Osmanovic Thunström, Almira/Steingrimsson, Steinn (2022): Can GPT-3 write an academic paper on itself, with minimal human input? hal.science/hal-03701250.
- GI – Gesellschaft für Informatik (2018): Die ethischen Leitlinien der Gesellschaft für Informatik e.V. Berlin/Bonn: GI. [gi.de/ueber-uns/organisation/unsere-ethischen-leitlinien](https://www.gi.de/ueber-uns/organisation/unsere-ethischen-leitlinien).
- Gmyrek, Paweł/Berg, Janine/Bescond, David (2023): Generative AI and Jobs – A global analysis of potential effects on job quantity and quality. ilo.org/wcmsp5/groups/public/---dgreports/---inst/documents/publication/wcms_890761.pdf.
- Hansen, Lukas (2018): 50 Millionen Nutzer. Datenaffäre setzt Facebook unter Druck. Deutsche Welle. [p.dw.com/p/2uaKl](https://www.dw.com/p/2uaKl).
- Hartong, Sigrid (2019): Learning Analytics und Big Data in der Bildung – Zur notwendigen Entwicklung eines datenpolitischen Alternativprogramms. Frankfurt am Main: GEW. [hsu-hh.de/ggb/wp-content/uploads/sites/679/2019/11/2019Hartong_Learning-analytics_GEW.pdf](https://www.hsu-hh.de/ggb/wp-content/uploads/sites/679/2019/11/2019Hartong_Learning-analytics_GEW.pdf).
- Heesen, Jessica et al. (2020): Zertifizierung von KI-Systemen – Kompass für die Entwicklung und Anwendung vertrauenswürdiger KI-Systeme. München: acatech. [acatech.de/publikation/zertifizierung-von-ki-systemen-kompass-fuer-die-entwicklung-und-anwendung-vertrauenswuerdiger-ki-systeme/download-pdf?lang=de](https://www.acatech.de/publikation/zertifizierung-von-ki-systemen-kompass-fuer-die-entwicklung-und-anwendung-vertrauenswuerdiger-ki-systeme/download-pdf?lang=de).
- Heitsch, Ernst (1993): Platon Werke III 4 Phaidros. Göttingen: Vandenhoeck & Ruprecht.
- Herzig, Bardo (2001): Medienerziehung und informatische Bildung – Ein (semiotischer) Beitrag zur integrativen Medienbildungstheorie. In: Herzig, Bardo (Hrsg.): Medien machen Schule. Grundlagen, Konzepte und Erfahrungen zur Medienbildung. Bad Heilbrunn: Klinkhardt, S. 129–164.
- Hoche, Johann Gottfried (1981[1794]): Vertraute Briefe über die jetzige abenteuerliche Lesesucht und über den Einfluss derselben auf die Verminderung des häuslichen und öffentlichen Glücks (Original: Hannover: Ritscher). München: Kraus.
- Hurrelmann, Klaus (2006): Einführung in die Sozialisationstheorie. Weinheim: Beltz.
- infas – Institut für angewandte Sozialwissenschaft (2018): Das Vermächtnis – die Welt, in der wir leben wollen. [infas.de/studien/das-vermaechtnis-die-welt-die-wir-erleben-wollen](https://www.infas.de/studien/das-vermaechtnis-die-welt-die-wir-erleben-wollen).
- Jacobs, Luisa (2023): Wie man mit KI wertvolle Arbeitszeit spart. Welche Tricks man kennen muss, um pünktlich Feierabend zu machen. [zeit.de/arbeit/2023-08/ki-arbeit-praesentation-mail-excel-tipps](https://www.zeit.de/arbeit/2023-08/ki-arbeit-praesentation-mail-excel-tipps).
- Kammerl, Rudolf/Kramer, Michaela/Müller, Jane/Potzel, Katrin/Tischer, Moritz/Wartberg, Lutz (2023): Dark Patterns und Digital Nudging in Social Media – wie erschweren Plattformen ein selbstbestimmtes Medienhandeln? Baden-Baden: Nomos. [blm.de/files/pdf2/blm-schriftenreihe_110.pdf](https://www.blm.de/files/pdf2/blm-schriftenreihe_110.pdf).
- Keil, Reinhard (1994): Das Gedächtnis lernt laufen – Vom Kerbholz zur virtuellen Wirklichkeit. In: Faßler, Manfred/Halbach: Wulf R. (Hrsg.): Cyberspace. Gemeinschaften, Virtuelle Kolonien, Öffentlichkeit. München: Wilh. Fink.
- Keil, Reinhard (2003): Technik als Denkzeug – Lerngewebe und Bildungsinfrastrukturen. In: Keil-Slawik, Reinhard/Kerres, Michael (Hrsg.): Wirkungen und Wirksamkeit Neuer Medien in der Bildung. Münster: Waxmann, S. 31–44.
- KMK (2016): Strategie der Kultusministerkonferenz „Bildung in der digitalen Welt“. Berlin: KMK, [kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_12_08-Bildung-in-der-digitalen-Welt.pdf](https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_12_08-Bildung-in-der-digitalen-Welt.pdf).
- Knaus, Thomas (2009): Kommunigrafie – Eine empirische Studie zur Bedeutung von Text und Bild in der digitalen Kommunikation. München: kopaed.

- Knaus, Thomas (2013): Technik stört! – Lernen mit digitalen Medien in interaktionistisch-konstruktivistischer Perspektive. In: Knaus, Thomas/Engel, Olga (Hrsg.): *fraMediale – digitale Medien in Bildungseinrichtungen* (Bd. 3). München: kopaed, S. 21–60. doi.org/10.25656/01:11687.
- Knaus, Thomas (2017): Verstehen – Vernetzen – Verantworten. Warum Medienbildung und informatische Bildung uns alle angehen und wir sie gemeinsam weiterentwickeln sollten. In: Diethelm, Ira (Hrsg.): *Informatische Bildung zum Verstehen und Gestalten der digitalen Welt (INFOS 2017)*, LNI – Lecture Notes in Informatics, P-274, Bonn: GI – Gesellschaft für Informatik, S. 31–48. doi.org/10.25656/01:14862.
- Knaus, Thomas (2018): Technikkritik und Selbstverantwortung – Plädoyer für ein erweitertes Medienkritikverständnis. In: Niesyto, Horst/Moser, Heinz (Hrsg.): *Medienkritik im digitalen Zeitalter* (Bd. 11 der Schriftenreihe Medienpädagogik interdisziplinär). München: kopaed, S. 91–107. doi.org/10.25656/01:17190.
- Knaus, Thomas (2019): Euphorie und Zukunftsangst. Zur subjekt- und kulturtheoretischen Sicht auf digitale Wandlungsprozesse (Vortrag im Rahmen der Ringvorlesung Technikfolgenabschätzung des IAO und des IAT der Universität Stuttgart). In: *Ludwigsburger Beiträge Zur Medienpädagogik – LBzM*, 20, S. 1–3. doi.org/10.21240/lbzm/20/14.
- Knaus, Thomas (2020a): Technology criticism and data literacy – The case for an augmented understanding of media literacy. In: *Journal of Media Literacy Education – JMLE*, 12 (3), pp. 6–16. doi.org/10.23860/JMLE-2020-12-3-2.
- Knaus, Thomas (2020b): Von medialen und technischen Handlungspotentialen, Interfaces und anderen Schnittstellen. In: Knaus, Thomas/Merz, Olga (Hrsg.): *Schnittstellen und Interfaces – Digitaler Wandel in Bildungseinrichtungen* (Bd. 7). München: kopaed, S. 15–72. doi.org/10.25656/01:18452.
- Knaus, Thomas (2022): It's the (inter)action, stupid! Einsam? Gemeinsam! ...über gute Lehre nachdenken. In: Knaus, Thomas/Junge, Thorsten/Merz, Olga (Hrsg.): *Lehren aus der Lehre in Zeiten von Corona – Mediendidaktische Impulse für Schulen und Hochschulen*. München: kopaed, S. 263–275. doi.org/10.25656/01:27484.
- Knaus, Thomas/Engel, Olga (2015): (Auch) auf das Werkzeug kommt es an – Technikhistorische und techniktheoretische Annäherungen an den Werkzeugbegriff in der Medienpädagogik. In: Knaus, Thomas/Engel, Olga (Hrsg.): *fraMediale – digitale Medien in Bildungseinrichtungen* (Bd. 4), München: kopaed, S. 15–57. doi.org/10.25656/01:11678.
- Knaus, Thomas/Engel, Olga (2016): *Wi(e)derstände – Digitaler Wandel in Bildungseinrichtungen* (Bd. 5). München: kopaed.
- Knaus, Thomas/Engel, Olga (2018): *Spannungen und Potentiale – Digitaler Wandel in Bildungseinrichtungen* (Bd. 6). München: kopaed.
- Knaus, Thomas/Junge, Thorsten/Merz, Olga (2022): *Lehren aus der Lehre in Zeiten von Corona. Mediendidaktische Impulse für Schulen und Hochschulen*. München: kopaed.
- Knaus, Thomas/Merz, Olga/Junge, Thorsten (2023): Ist das Kunst... oder kann das die KI? Zum Verhältnis von menschlicher und künstlicher Kreativität, Cfp 24/2024, S. 1–8. medienpaed-ludwigsburg.de/public/journals/1/cfp/LBzM-Heft24_Cfp-KI_Kreativitaet.pdf.
- Knaus, Thomas/Schmidt, Jennifer (2020): Medienpädagogisches Making. In: *Medienimpulse*, 58 (4), S. 1–25. journals.univie.ac.at/index.php/mp/article/view/4322.
- Knaus, Thomas/Schmidt, Jennifer/Merz, Olga (2023): Aktive Medienarbeit als Vorbild – Handlungsorientierte Ansätze zur Förderung einer um digitaltechnische Dimensionen erweiterten Medienbildung. In: *medien+erziehung – merz*, 67 (03), S. 42–49. doi.org/10.25656/01:26533.
- Knaus, Thomas/Tulodziecki, Gerhard (2023): Thomas Knaus im Gespräch mit Gerhard Tulodziecki. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 2023/23, S. 1–23. doi.org/10.21240/lbzm/23/22.

- Kommer, Sven/Missomelius, Petra/Knaus, Thomas/Büsch, Andreas (2016): Stellungnahme zum KMK Strategiepapier „Bildung in der digitalen Welt“. keine-bildung-ohne-medien.de/stellungnahme/mkmstrategiepapier.
- Krämer, Sybille (1988): *Symbolische Maschinen. Die Idee der Formalisierung in geschichtlichem Abriß*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Krempl, Stefan (2023): Arbeitspause für mächtige Sprachmodelle – Internationale Technikexperten fordern eine Arbeitspause an mächtigen Sprachmodellen. heise.de/news/Moratorium-fuer-grosse-KI-Modelle-ChatGPT-Co-koennen-Leben-retten-8456050.html.
- Kümmel, Albert/Scholz, Leander/Schumacher, Eckhard (2004): *Einführung in die Geschichte der Medien*. Paderborn: Wilh. Fink.
- Langville, Amy N./Meyer, Carl D. (2006): *Google's PageRank and beyond – The science of search engine rankings*. Princeton (NJ): Princeton University Press.
- Lankau, Ralf (2022): *Kein Mensch lernt digital – Über den sinnvollen Einsatz neuer Medien im Unterricht*. Weinheim: Beltz.
- Marchand, André (2023): Künstliche Intelligenz und die Abschlussarbeit. ChatGPT hat längst Einzug gehalten ins Leben der Studierenden. Dadurch wird sich die Kompetenzüberprüfung ändern müssen. forschung-und-lehre.de/lehre/kuenstliche-intelligenz-und-die-abschlussarbeit-5819.
- Mayrberger, Kerstin (2010): *Neues Lernen mit neuen Medien 2.0 – Potentiale und Grenzen von Personal Learning Environments (PLE) für Schule und Hochschule*, In: Bauer, Petra/Hoffmann, Hannah/Mayrberger, Kerstin (2010): *Fokus Medienpädagogik – Aktuelle Forschungs- und Handlungsfelder*. München: kopaed, S. 280–298.
- McLuhan, Marshall (1964): *Understanding Media – The Extensions of Man*. New York: McGraw-Hill.
- Meister, Dorothee M. (2008): *Schule und Medien*. In: Helsper, Werner/Böhme, Jeanette (Hrsg.): *Handbuch der Schulforschung*. Wiesbaden: Springer, S. 513–529.
- Meyer-Drawe, Käte (1995): *Von der Marionette bis zum autopoietischen System – Maschinenbilder in der Pädagogik*. In: *Vierteljahrsschrift für wissenschaftliche Pädagogik*, 71 (4), S. 358–373.
- MPFS – Medienpädagogischer Forschungsverbund Südwest (2021a): *JIM-Studie 2021*. mpfs.de/fileadmin/files/Studien/JIM/2021/JIM-Studie_2021_barrierefrei_230418_neu.pdf.
- MPFS – Medienpädagogischer Forschungsverbund Südwest (2021b): *SIM-Studie 2021*. mpfs.de/fileadmin/files/Studien/SIM/2021/Web_SIM-Studie2021_final_barrierefrei.pdf.
- MPFS – Medienpädagogischer Forschungsverbund Südwest (2022): *JIM-Studie 2022*. mpfs.de/fileadmin/files/Studien/JIM/2021/JIM-Studie_2021_barrierefrei_230418_neu.pdf.
- Mullin, Benjamin/Grant, Nico (2023): *Google tests A.I. tool that is able to write news articles*. nytimes.com/2023/07/19/business/google-artificial-intelligence-news-articles.html.
- Nagel, Marie-Theres/Zlatkin-Troitschanskaia, Olga/Fischer, Jennifer (2023): *Validation of newly developed tasks for the assessment of generic Critical Online Reasoning (COR) of university students and graduates*. doi.org/10.3389/feduc.2022.914857.
- Nehlsen, Johannes/Fleck, Tilmann (2023): *Ist ChatGTP ein zulässiges Hilfsmittel in Prüfungen? Der Chatbot ChatGPT wirft an Unis prüfungsrechtliche Fragen auf*. In: *Forschung&Lehre*, 4/23, S. 262–264.
- Nilsson, Nils John (2010): *The quest for artificial intelligence – A history of ideas and achievements*, Cambridge: CUP.
- Pfemfert, Franz (1992 [1911]): *Kino als Erzieher*. In: Schweinitz, Jörg (Hrsg.): *Prolog vor dem Film. Nachdenken über ein neues Medium 1909 bis 1914*. Leipzig: Reclam, S. 165–169.
- Rath, Matthias O./Maisenholder, Patrick (2021): *Digitalisierung als Capability und Fairness. Ausblicke auf eine Postcorona-Lehre*. In: *Ludwigsburger Beiträge Zur Medienpädagogik – LBzM*, 21, S. 1–20. doi.org/10.21240/lbzm/21/02.

- Reich, Kersten (2008): *Konstruktivistische Didaktik*. Weinheim: Beltz.
- Reinholz, Daniel L./Stone-Johnstone, Amelia/Shah, Niral (2020): Walking the walk – Using classroom analytics to support instructors to address implicit bias in teaching. In: *The International Journal for Academic Development* 25, 259–272. doi.org/10.1080/1360144X.2019.1692211.
- Ritter, Helge/Martinetz, Thomas/Schulten, Klaus (1992): *Neural computation and self-organizing maps – An introduction*. Boston, MA: Addison-Wesley.
- Rössler, D. (1983). Die Sicherheit der Technik als öffentliches Problem. In: Hartwig, Sylvius (Hrsg.): *Große technische Gefahrenpotentiale*. Heidelberg: Springer.
- Sachse, Maximilian/Thielen, Johannes (2023): Die große Bonanza mit Künstlicher Intelligenz. faz.net/aktuell/wirtschaft/schneller-schlau/chatgpt-keine-internetanwendung-wuchs-je-so-schnell-18799756.html.
- Salden, Peter/Lordick, Nadine/Wiethoff, Maike (2023): KI-Basierte Schreibwerkzeuge in der Hochschule. In: Salden, Peter/Leschke, Jonas (Hrsg.): *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung*. hss-opus.ub.ruhr-uni-bochum.de/opus4/frontdoor/index/index/docId/9734.
- Schelhowe, Heidi (1997): *Das Medium aus der Maschine – Zur Metamorphose des Computers*. Frankfurt am Main: Campus.
- Schmidt, Siegfried J. (2000): *Kalte Faszination. Medien – Kultur – Wissenschaft in der Mediengesellschaft*. Weilerswist: Velbrück.
- Scholz, Leander (2004): Die Industria des Buchdrucks. In: Kümmel, Albert/Scholz, Leander/Schumacher, Eckhard (Hrsg.): *Einführung in die Geschichte der Medien*. Paderborn: Willh. Fink, S. 11–33.
- Schorb, Bernd (1982): Bildungs- und Schulfernsehen. In Kreuzer, Helmut (Hrsg.): *Sachwörterbuch des Fernsehens*. Göttingen: Vandenhoeck & Ruprecht, S. 25–31.
- Schorb, Bernd (1994): Bildungsfernsehen. In Ludes, P/Schumacher, H/Zimmermann, P (Hrsg.): *Informations und Dokumentarsendungen – Geschichte des Fernsehens in der Bundesrepublik Deutschland* (Bd. 3). München: Fink, S. 203–212.
- Schubert, Christian (2023): Italien sperrt Chat GPT. faz.net/aktuell/italien-sperrt-chat-gpt-18791191.html.
- Seemann, Michael (2021): *Die Macht der Plattformen – Politik in Zeiten der Internetgiganten*. Berlin: Aufbau Verlag.
- Shani, Inbal (2023): The Developer Wishlist – Developers think AI increases productivity and prevents burnout. Github. github.blog/2023-06-13-survey-reveals-ais-impact-on-the-developer-experience.
- Shneiderman, Ben (2020): Human-centered artificial intelligence – Reliable, safe & trustworthy. In: *International Journal of Human-Computer Interaction*, 36 (6), pp. 495–504.
- Sickert, Teresa (2016): Vom Hipster-Mädchen zum Hitler-Bot. spiegel.de/netzwelt/web/microsoft-twitter-bot-tay-vom-hipstermaedchen-zum-hitlerbot-a-1084038.html.
- Siebert, Horst (2005): *Pädagogischer Konstruktivismus*. Weinheim: Beltz.
- Siepmann, Dirk (2023): Auswirkungen von KI auf die Textproduktion in der Wissenschaft. In: *Forschung & Lehre*, 7/23. forschung-und-lehre.de/zeitfragen/welche-auswirkungen-kis-auf-die-textproduktion-in-der-wissenschaft-haben-5740.
- Singhal, Karan/Azizi, Shekoofeh/Tu, Tao et al. (2023): Large language models encode clinical knowledge. In: *Nature* 620, pp. 172–180. doi.org/10.1038/s41586-023-06291-2.
- Skinner, Burrhus Frederic (1954): The science of learning and the art of teaching. In: *Harvard Educational Review*, 2 (2), pp. 86–97.
- Spannagel, Christian (2023a): ChatGPT und die Zukunft des Lernens: Evolution statt Revolution. hochschulforumdigitalisierung.de/de/print/blog/chatgpt-und-die-zukunft-des-lernens-evolution-statt-revolution.

- Spannagel, Christian (2023b): Rules for Tools. csp.uber.space/phhd/rulesfortools.pdf.
- Stokel-Walker, Chris (2023): ChatGPT listed as author on research papers – Many scientists disapprove. In: *Nature* 613, pp. 620–621. doi.org/10.1038/d41586-023-00107-z.
- Swertz, Christian (2017): Medienkompetenz und digitale Bildung aus medienpädagogischer Perspektive. In: *MedienImpulse*, 55 (1). doi.org/10.25656/01:15549.
- Tickle, Alan B./Andrews, Robert/Golea, Mostefa/Diederich, Joachim (1998): The truth will come to light – Directions and challenges in extracting the knowledge embedded within trained artificial neural networks. In: *Transactions on Neural Networks*, 9 (6), S. 1057–1068. doi.org/10.1109/72.728352.
- Tulodziecki, Gerhard (2023a): Individuelles Handeln und Gemeinwohl. Eine interdisziplinäre Handlungstheorie im Kontext von Freiheit, Verantwortung und künstlicher Intelligenz. Bielefeld: transcript.
- Tulodziecki, Gerhard (2023b): Medienkompetenz und Handlungstheorie. Zu Schwächen und Desiderata medienpädagogischer Überlegungen aus handlungstheoretischer Sicht. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 23/2023, S. 1–15. doi.org/10.21240/lbzm/23/02.
- Volckart, Benedikt (2022): Digitalisierung und Schule – Vom Beitrag der Medienpädagogik für die Schulentwicklung. Interview für das German Chapter (GChACM) der ACM, Q3/2022. germany.acm.org/2022/08/09/digitalisierung-und-schule-teil-2-vom-beitrag-der-medienpaedagogik-fuer-die-schulentwicklung.
- Vuorikari, Riina/Kluzer, Stefano/Punie, Yves (2022): DigComp 2.2: The Digital Competence Framework for Citizens – With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022. doi.org/10.2760/490274,JRC128415.
- Weber-Stein, Florian/Heidenreich, Felix (2023): Digitalisierung als Pharmakon. Medienbildung unter Bedingungen der Post-Souveränität. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, Ausgabe 23/2023, S. 1–15. doi.org/10.21240/lbzm/23/05.
- Weck Andreas (2023): Deep Fakes als Propaganda-Waffe – Giffey spricht mit angeblichem Klitschko. t3n.de/news/deep-fakes-propaganda-waffe-1481796.
- Weiß, Eva-Maria (2023): Der KI-Papst im Daunenmantel sollte eine Warnung sein – eine Analyse. heise.de/hintergrund/Der-KI-Papst-in-Daunenmantel-sollte-eine-Warnung-sein-8146920.html.
- Weizenbaum, Joseph (1966): ELIZA – A computer program for the study of natural language communication between man and machine. In: *Communications of the ACM*. dl.acm.org/doi/10.1145/365153.365168.
- Welskopp, Thomas (2010): *Amerikas große Ernüchterung – Eine Kulturgeschichte der Prohibition*. Paderborn: Schöningh.
- Weßels, Doris (2021): Verführerische Werkzeuge – Plagiate und KI-gestützte Textproduktion an Hochschulen? In: *Forschung&Lehre*, 12/2021, S. 1018–1019. forschung-und-lehre.de/zeitfragen/ki-gestuetzte-textproduktion-an-hochschulen-4292.
- Weßels, Doris/Gottschalk, Ole (2023): Hochschullehre unter dem Einfluss des KI-gestützten Schreibens – Wie soll die Hochschullehre mit Texten umgehen, die von einer KI verfasst wurden? hochschulforumdigitalisierung.de/de/print/blog/Hochschullehre-KI-gestuetztes-Schreiben.
- Wessler, Hartmut (1997): *Perspektiven der Medienkritik – Die gesellschaftliche Auseinandersetzung mit öffentlicher Kommunikation in der Mediengesellschaft*. Opladen: Westdeutscher Verlag.
- Wittenhorst, Tilman (2023): ChatGPT erfindet Gerichtsurteile – Ein US-Rechtsanwalt befragt ChatGPT nach Präzedenzfällen und übernimmt die Antwort ungeprüft. heise.de/news/ChatGPT-erfindet-Gerichtsurteile-US-Anwalt-faellt-darauf-herein-9068180.html.
- Wolfram, Stephen (2023): What is ChatGPT doing... and why does it work? writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work.

Informationen zu den Autor*innen



Thomas Knaus ist Professor für *Erziehungswissenschaft SP Medienpädagogik* und Leiter der Abt. Medienpädagogik der PH Ludwigsburg, Wiss. Direktor des FTzM und Honorarprofessor für *Bildungsinformatik* am Fb Informatik der Frankfurt UAS; zuvor war er als Universitätsprofessor für Allg. Erziehungswissenschaft in Erlangen-Nürnberg sowie u. a. an der Universität Wien als Visiting Professor tätig; seine Arbeits- und Forschungsschwerpunkte sind u. a. der digitale Wandel in Bildungseinrichtungen; er engagiert sich u. a. im Lenkungskreis der Initiative KBoM, im erweiterten Vorstand der GMK, der Fachgruppe Schule sowie als Sprecher der Fachgruppe *Qualitative Forschung*; www.thomas-knaus.de.



ChatGPT ist ein fortschrittliches KI-Sprachmodell, entwickelt von *OpenAI*. Es basiert auf der leistungsstarken GPT-3.5-Architektur und verfügt über ein umfangreiches Wissen in verschiedenen Bereichen. ChatGPT wurde darauf trainiert, komplexe Fragen zu beantworten, informative und kreative Antworten zu generieren und Nutzer[*inne]n bei der Lösung von Problemen und der Bereitstellung von Informationen zu helfen. Es ist eine innovative KI-Lösung, die darauf abzielt, das menschliche Schreib- und Kommunikationserlebnis zu verbessern.³⁶

Zitationshinweis:

Knaus, Thomas et AI (2023): Künstliche Intelligenz und Bildung: Was sollen wir wissen? Was können wir tun? Was dürfen wir hoffen? Und was ist diese KI? Ein kollaborativer Aufklärungsversuch. In: *Ludwigsburger Beiträge zur Medienpädagogik – LBzM*, 23/2023. S. 1–42. doi.org/10.21240/lbzm/23/19

³⁶ Selbstvorstellung der Co-Autor*in ChatGPT [Prompt: bitte stelle Dich für eine Autor*in-Biografie in drei/vier Sätzen kurz vor, nicht in der Ich-Form].