
Themenheft 58: Spannungsfeld der digitalen Kompetenz.

Herausgegeben von Miriam Mulders, Kristian Träg, Tatjana Steinhaus und Anne Vonarx

Künstliche Intelligenz im Kontext von Kompetenzen, Prüfungen und Lehr-Lern-Methoden

Alte und neue Gestaltungsfragen

Maria Klar¹  und Johannes Schleiss² 

¹ Universität Duisburg-Essen

² Otto-von-Guericke-Universität Magdeburg

* Der vorliegende Beitrag wurde in geteilter Erstautor:innenschaft verfasst.

Zusammenfassung

Der Beitrag untersucht Auswirkungen und Fragestellungen von generativen Sprachmodellen, einer Form der Künstlichen Intelligenz (KI), in den Bereichen von Kompetenzen, Prüfungen und Lehr-Lern-Szenarien. Wir diskutieren, inwiefern KI neue Kompetenzen erfordert, Kompetenzen an Relevanz verlieren und Kompetenzen verloren gehen. Die teilweise veränderten Kompetenzen wirken auch auf die Prüfungsgestaltung: Es stellt sich die Frage, wo KI aus Prüfungen ausgeschlossen und wo sie bewusst eingeschlossen werden sollte. Beide Entscheidungen haben Einfluss auf die Zuverlässigkeit und Alltagsnähe der Prüfungen. KI wirkt hier als Verstärker bestehender Fragen und Problemstellungen. Im Bereich der Lehr-Lern-Szenarien eröffnet KI Lernenden die Möglichkeit, Lernmedien selbst zu adaptieren. Darüber hinaus kann KI als weiterer Akteur in Lehr-Lern-Settings hinzukommen und es stellen sich Fragen, wie dies die soziale Situation des Lernens verändern kann. Im Sinne des Constructive Alignment argumentiert dieser Beitrag, dass KI nicht isoliert, sondern didaktisch verschränkt in Kompetenzen, Prüfungen und Lehr-Lern-Szenarien integriert werden muss. Für die Diskussion und Exploration der sich daraus ergebenden Fragen brauchen Bildungsinstitutionen Offenheit und Experimentierräume.

Artificial Intelligence in the Context of Competencies, Assessment and Instructional Design. Old and New Questions of Practices and Design

Abstract

This article examines the impacts and issues of generative language models, a form of artificial intelligence (AI), in the areas of competencies, assessments, and instructional design. We discuss to what extent AI requires new competencies, makes competencies lose relevance, and leads to the loss of competencies. Changes in competencies also affect the design of assessment scenarios: The question arises as to where AI should be excluded from assessments and where it should be deliberately included. Both decisions have an impact on the reliability and practical relevance of the assessment. Here, AI acts as an amplifier of existing questions and issues. In the area of instructional design, AI offers to learners the possibility to adapt learning media to their own needs. Furthermore, AI can participate as an additional actor in instructional settings, raising questions about how this can change the social situation of learning. Following the view of constructive alignment, this article argues that AI should not be discussed in an isolated way, but integrated into competencies, assessments and instructional designs in a holistic way. Educational institutions need openness and room for experimentation to discuss and explore the resulting questions.

1. Einleitung

Künstliche Intelligenz (KI) und ihre Bedeutung für die Bildung sind in aller Munde. ChatGPT ist die am schnellsten wachsende User-Anwendung der Geschichte (Hu und Hu 2023), und die rasante technische Entwicklung von generativen Modellen sorgt für immer neue Schlagzeilen. Besonders die Bildung wird immer wieder als Bereich genannt, der sich grundlegend durch generative Sprachmodelle (Large Language Models, LLMs) verändern soll (z. B. Ebbinghaus 2023). Auch innerhalb der Bildungscommunity ist das Interesse gross: Die Anzahl von Vorträgen, Diskussionsrunden und Gelegenheiten zum Ideenaustausch bleibt auf hohem Niveau und erreicht häufig sehr hohe Teilnehmendenzahlen.¹ Doch welche neuen Fragen stellen sich tatsächlich durch den Einzug von grossen Sprachmodellen in Bildungslandschaften? Wo bleiben alte Fragen bestehen? Wie wollen wir den Diskurs und die Entwicklung gestalten?

¹ Der vorliegende Beitrag ist die Fortführung eines einstündigen Workshops im Rahmen der Tagung «Junges Forum Medien und Hochschulentwicklung (JFMH)» im Juli 2023 mit dem Titel «Unterwegs im KI-Hypetrain». Orientierend an der Zug-Metapher durchliefen die 30 Teilnehmenden, primär Forschende, die Stationen von Kompetenzen, Prüfungen und Lehr- und Lernaktivitäten, wobei sie an jeder Station auf Basis eines kurzen Impulses innerhalb ihrer «Zugabteile» kurz über die Auswirkungen diskutierten.

Bereits in einer früheren Hochphase von Sprachmodellen wurden Auswirkungen unterschiedlicher Art diskutiert (z. B. Weizenbaum 1976). Heutige Erkenntnisse bauen auf mehreren Jahrzehnten Forschung in diesem Bereich auf (Bond et al. 2024; Zawacki-Richter et al. 2019). Gleichzeitig ist zu beachten, dass die Nutzung von KI in einem Bereich von Bildungsangeboten auch notwendige Veränderungen in anderen Bereichen nach sich ziehen kann. In diesem Beitrag betrachten wir daher die Integration von KI, insbesondere von generativen Sprachmodellen in den Bereichen der *Kompetenzen*, der *Prüfungen* sowie der *Lehr- und Lernaktivitäten*. Diese Bereiche folgen der Aufgliederung des Constructive Alignments (Biggs 1996), wonach Lehre die Aufgabe hat, Lehr- und Lernziele zu operationalisieren und nach diesen Zielen auch die Prüfungsszenarien und das Lernangebot auszurichten. In diesem Sinne lässt sich KI nicht isoliert nur in einen Bereich der Lehre integrieren. Wie im Folgenden dargelegt wird, ist es beispielsweise nicht sinnvoll, einerseits Fähigkeiten und Wissen mit KI-Bezug in die Curricula aufzunehmen, ohne zu prüfen, ob Prüfungsformate und Lernangebote entsprechend angepasst werden müssen.

Ausgangspunkt dieses Beitrags ist daher die Frage nach veränderter Kompetenz, also Veränderungen in Konglomeraten aus Wissen, Fähigkeiten und Einstellungen (Weinert 2001). Von dort aus fokussiert der Beitrag den Diskurs rund um Prüfungen und zeigt auf, dass KI hier zu einer Verstärkung bisheriger Fragen und Problemstellungen führt. Schliesslich geht er auf KI im Kontext von Lehr- und Lernaktivitäten ein und beschreibt, wie (generative) KI als neues Werkzeug und neuer Akteur wirkt. Abschliessend werden die wichtigsten Erkenntnisse zusammengefasst und diskutiert sowie auf weitere Fragen, die es nun zu diskutieren gilt, verwiesen.

2. Kompetenzen

Allgemein verfügbare generative KI befeuert den Diskurs rund um Kompetenzen. Es stellen sich die Fragen, ob neue Kompetenzen im Umgang mit KI benötigt werden («KI-Kompetenzen»), ob andere Kompetenzen an Bedeutung verlieren und ob wir Kompetenzen verlernen. Angesichts von Ungewissheiten zeigt sich ein Wunsch nach Orientierung, beispielsweise in der Formulierung von «Zukunftskompetenzen» (z. B. Ehlers 2020). Jedoch ist fraglich, inwiefern dieses Versprechen, sich auf eine ungewisse Zukunft vorbereiten zu können, einlösbar ist (vgl. Engel und Kerres 2023).

Eine zentrale Frage ist, ob «KI-Kompetenzen» eigenständige Kompetenzen sind oder bestehende Kompetenzen durchdringen. Während oft die Rede von «digitalen Kompetenzen» ist, mit der Implikation, dass es sich hierbei um ein eigenständiges Set an additiven Kompetenzen handelt, argumentiert Kerres (2017), dass digitale Technologien bestehende Kulturtechniken durchdringen und ein separates Lehren und Lernen von «digitalen Kompetenzen» nicht sinnvoll ist.

Ein Versuch der Definition von KI-Kompetenzen erfolgt durch den Diskurs zu Data- und AI Literacy (Long und Magerko 2020; Schüller et al. 2023) und wird auch durch immer neue Rahmenwerke von Organisationen – wie der UNESCO, die aktuell an einem Kompetenzrahmen für Lehrende und Studierende arbeitet, oder der EU mit dem DigComp Framework (Vuorikari, Kluzer, und Punie 2022) – gefördert. Im Zentrum steht die Annahme, dass es neue Kompetenzen im Umgang mit Daten, vor allem mit KI-Systemen braucht. KI-Kompetenzen sind hierbei als Umgang mit KI-Technologie zu interpretieren. Dass Daten als Grundlage und KI-Systeme als Werkzeuge eine immer grössere Bedeutung in allen Disziplinen haben, führt zu einer Veränderung der Anforderungen in verschiedenen Arbeitsbereichen (Schleiss et al. 2023). Ein Beispiel ist die Fähigkeit, effektiv Prompts zu formulieren und Strategien zum Arbeiten mit generativen Tools und Werkzeugen zu entwickeln (Bozkurt und Sharma 2023). Dies ist vergleichbar mit der Suchkompetenz, die mit Beginn des Internets aufkam und oft als eine Dimension der *Informationskompetenz* geführt wird (Leber et al. 2023).

Durch generative Sprachmodelle und entsprechende Tools ändert sich auch die Art und Weise, wie auf Wissen zugegriffen wird. Erste nicht-repräsentative Studien zur Nutzung von generativer KI an deutschen Hochschulen zeigen, dass 63% der Studierenden bereits generative Modelle im Studium einsetzen und die Haupteinsatzgebiete das Klären von (fach-)spezifischen Fragestellungen, Forschungs- und Literaturarbeit, Übersetzungen sowie die Problemlösung und Entscheidungsfindung sind (von Garrel und Mayer 2023). Jedoch bleibt die Frage, inwiefern KI-Technologie hier zu gänzlich neuen Wissenspraktiken führt, aus denen sich neue Kompetenzen ableiten lassen, oder ob bestehende Praktiken um die Aspekte dieser Technologien erweitert werden.

Kalz (2023) zeigt hinsichtlich der Formulierung von *Zukunftskompetenzen* einige Problemfelder auf, die auch für die Diskussion um AI Literacy relevant sind: Es braucht eine Anbindung an bisherige Kompetenzmodelle, statt neue Listen von (KI-) Kompetenzen aufzustellen, deren Auswahl und Priorisierung neu begründet werden muss. Weiterhin zeigt sich die Messung solcher Kompetenzen als Herausforderung und es braucht empirische Forschung dazu, welche Effekte ein Zuwachs an KI-Kompetenzen in Performanzsituationen erreicht. Analog zur Digitalkompetenz stellt sich auch bei *AI Literacy* die Frage, ob es sich nicht vielmehr um eine Pluralität von Kompetenzen handelt, die darüber hinaus nur in der fachlichen Verankerung lehr- und lernbar sind (Kerres 2023).

Zusammenfassend lässt sich ein Problem der Begrifflichkeiten feststellen. KI-Kompetenzen beinhalten bestehende Kompetenzen und sind nicht grundlegend neu. Die Art der Nutzung von Daten und KI führt aber zu neuen Wissenspraktiken und erhöht die Wichtigkeit eines grundlegenden Verständnisses über die Funktionsweise von KI-Algorithmen und Daten.

Auch wenn es gelänge, KI-Kompetenzen zu formulieren, die auch in Zukunft Bestand haben sollen, bleiben Unsicherheiten: Dadurch, dass die KI-basierten Werkzeuge gewisse Dinge schneller und zuverlässiger als Menschen erledigen können, verändern sich Bedeutung und Wichtigkeit dieser Kompetenzen. Beispielsweise dauerte es mitunter Jahre, um zu entschlüsseln, wie sich ein einzelnes Protein faltet. Ein KI-Modell namens AlphaFold hat dies nun in einem Bruchteil der Zeit für einen Grossteil der Proteine gelöst (Jumper et al. 2021). Wie gehen wir damit um, dass Kompetenzen, die wir besitzen oder gerade lernen, möglicherweise aufgrund technischer Automatisierung irrelevant werden? Wie gehen wir mit den sozialen und gesellschaftlichen Dimensionen um, beispielsweise einem Verlust von Status oder einem Gefühl der Kränkung, wenn menschliche Arbeitsleistung nicht mehr konkurrenzfähig ist? Das *Upskilling* oder *Reskilling* könnte durch die dynamische technische Entwicklung erschwert werden (Hamilton, Wiliam, und Hattie 2023). Die Vorhersage, welche Tätigkeiten in naher Zukunft von Maschinen ausgeführt werden können, erweist sich als schwierig.

Weiterhin stellt sich die Frage nach dem Verlernen von Kompetenzen, auch *Deskilling* genannt. Durch die Nutzung von immer stärkeren Werkzeugen geben wir Verantwortung auf diese ab und könnten so die grundlegenden Fähigkeiten verlernen (Rafner et al. 2021; Reinmann 2023). Eine Erklärung bietet Forschung rund um *Cognitive Offloading*, die zeigt, dass ein temporäres Abladen auf technische Werkzeuge zwar kurzfristig zu besserer Performanz, langfristig aber zu schlechterer Behaltensleistung führt (Grinschgl, Papenmeier, und Meyerhoff 2021). Hamilton et al. (2023) sprechen gar von *de-education*, dem Verlust von Bildung dadurch, dass Menschen nicht mehr motiviert sind, sich Kompetenzen anzueignen angesichts einer übermenschlichen KI. Potenzielle Kompetenzverluste werden auch von Studierenden als Risiko gesehen (Gottschling, Seidl, und Vonhof 2023). Dabei ist das Auslagern von Prozessen an (digitale) Technologien an sich kein neuer Trend und kommt alltäglich beispielsweise in der Verwendung von automatischer Schreibkorrektur oder der Nutzung eines Taschenrechners vor. Eine zentrale Frage ist dabei, ob die Grundkompetenzen von Werkzeugen erst beherrscht werden müssen, bevor man diese einsetzt, wie es beispielsweise in der Didaktik der Mathematik unter dem Begriff *Reverse Scaffolding* diskutiert wird (Chase und Abrahamson 2015). Dabei soll die Grundkompetenz vor allem zu einer kritischen und reflektierten Nutzung sowie zu einer Verringerung der Abhängigkeit von digitaler Technologie führen (Reinmann 2023).

In diesem Zusammenhang ist des Weiteren eine Überschätzung der eigenen Fähigkeiten im KI-Bereich problematisch. Wie in unterschiedlichen Bereichen nachgewiesen wurde, neigen Personen mit wenig Vorwissen dazu, die eigenen Fähigkeiten zu überschätzen (Dunning 2011). Dies ist ebenfalls im Kontext von KI-Fähigkeiten

und verknüpften Fähigkeiten zu beobachten und kann zu einer Überschätzung der eigenen Fähigkeiten, verbunden mit einer Unterschätzung der Risiken führen (He, Kuiper, und Gadiraju 2023).

Angesichts dieser Diskussionen lässt sich festhalten, dass die Beschleunigung der technischen Entwicklung zu einer Veränderung und veränderten Gewichtung von Kompetenzen führt. Dabei geht es auf der einen Seite um eine (fachspezifische) Handlungskompetenz im Umgang mit KI-Systemen, andererseits auch um eine veränderte Bedeutung von menschlichen Fähigkeiten wie kritisches Denken, Empathie oder Vorstellungskraft in Abgrenzung zu maschinellen Fähigkeiten wie der Verarbeitung und Auswertung grosser Datenmengen. Diese Veränderung und Umgewichtung sollte auch dazu führen, die Auswirkungen der technischen Entwicklung auf die Prüfungsformate und die Methodik der Bildungsangebote zu diskutieren.

3. Prüfungen

Wenn wir nun annehmen, dass neue Fähigkeiten, neues Wissen oder neue Einstellungen in Bezug auf KI als Lehr- und Lernziele in (hoch)schulische Curricula aufgenommen werden, müssen diese Lehr- und Lernziele wiederum in der Gestaltung von Prüfungsszenarien Beachtung finden. Beispielsweise zeichnet sich ab, dass Praktiken des Schreibens sich mit generativen Sprachmodellen stärker verändern könnten als zuvor. Studienveranstaltungen, die beispielsweise unter anderem auf Schreibkompetenz abzielen, sollten den Lernenden die Chance bieten, den reflektierten und effektiven Umgang mit LLMs zu erproben. Dann allerdings sollten diese Werkzeuge nicht aus der Prüfungsleistung ausgeschlossen werden. Dies kann eine Herausforderung darstellen, da KI-basierte Werkzeuge womöglich einen (zu) grossen Teil der Leistung erbringen können, die in einer Prüfung gefragt ist.

Generative Sprachmodelle sind schon jetzt in der Lage, längere Textabschnitte von erläuternden, argumentativen oder reflexiven Texten zu produzieren. Das Kontextfenster, also die Menge an Input und Output, die durch ein LLM verarbeitet werden kann, wächst ständig und so lassen sich längere Texte mit zunehmend höherem Komplexitätsgrad maschinell erstellen. Es besteht die Möglichkeit, dass Lernende KI-generierte Texte als ihre Eigenleistung einreichen, ohne dass es eine Software gäbe, die diese verlässlich und diskriminierungsfrei erkennen kann (Khalil und Er 2023; Liang et al. 2023). Die klassische Hausarbeit ist somit heute kein verlässliches Instrument zur Zertifizierung von Lernleistungen mehr. Die Frage nach zulässigen Hilfsmitteln ist dabei nicht neu: In der Mathematik gibt es beispielsweise eine andauernde Debatte zum Umgang mit Taschenrechnern (Monaghan 2016) und an Hochschulen werden Open-Book-Klausuren eingesetzt (Brightwell, Daniel, und Stewart

2004). Doch generative Sprachmodelle stellen hier mehr als nur ein Hilfsmittel dar, sondern ein Werkzeug, um gängige Prüfungsleistungen auf einem passablen Niveau mit relativ wenig menschlichem Input zu generieren (Malinka et al. 2023).

Es gibt mindestens zwei – gleichermassen unvollständige – Antworten auf die Frage nach dem Umgang mit LLMs als Werkzeug in Prüfungszusammenhängen (Buck und Limburg 2023; Budde, Tobor, und Beyermann 2023): Man kann versuchen, KI-Werkzeuge aus Prüfungssituationen auszuschliessen und Prüfungen «betrugssicher» zu gestalten, also beispielsweise zu Präsenz-Klausuren zurückkehren. Letzteres würde stärker das Ziel der gerechten Zertifizierung von Leistungen fokussieren, wäre aber eine didaktische Einschränkung angesichts der begrenzten Möglichkeiten, komplexere Kompetenzzusammenhänge in diesem Format sichtbar zu machen. Alternativ kann man Sprachmodelle als neue Werkzeuge einschliessen und Prüfungsformate so komplex gestalten, dass Lernende herausgefordert sind, mit den generierten Ergebnissen zu arbeiten und über diese hinauszugehen.² Dies kann beispielsweise in Gruppenprüfungen, mündlichen Verteidigungen oder Projektprüfungen geschehen. Solche Prüfungsformate würden das Ziel des Kompetenzerwerbs und der Alltagsnähe von Prüfungen in den Fokus nehmen, können allerdings ressourcenintensiv sein.

Um den Korrekturaufwand, der durch komplexere Prüfungsformate entsteht, entlasten zu können, könnte wiederum KI hilfreich sein. Langjährige Forschung zur automatischen Bewertung von Prüfungsleistungen gibt es besonders im US-amerikanischen Raum, wo Essays und Hausarbeiten eine grosse Rolle spielen (Paiva, Leal, und Figueira 2022; Ramesh und Sanampudi 2022). Ironischerweise kann es nun vermehrt dazu kommen, dass KI-Systeme Essays bewerten, die von KI-Systemen geschrieben wurden.

Über einen reduzierten Aufwand bei Prüfungen hinaus kann KI dazu genutzt werden, Lernprozesse zu unterstützen. Unterstützung bei Korrekturen könnte Lehrenden mehr Freiraum geben, um nicht nur summative, sondern auch häufiger formative Rückmeldung, also Feedback im Verlauf des Lernprozesses zu geben (Zhai, He, und Krajcik 2022). Solch formatives Feedback wiederum könnte ebenfalls durch KI unterstützt werden (z. B. Benotti et al. 2018). Ein weiteres Beispiel ist die Entwicklung von Analysewerkzeugen, die kleinschrittig den Lernprozess evaluieren und Lernenden Feedback über den Prozess geben, oft unter dem Begriff *Learning Analytics* zusammengefasst (Ifenthaler und Drachslar 2020). Eine Herausforderung dieser Learning Analytics-Werkzeuge besteht jedoch noch darin, qualitativ tiefgehende Analysen vorzunehmen, die über Klickverhalten auf Lernplattformen und Auswertung von kurzen Tests hinausgehen (Stanja et al. 2023). Selbst in standardisierten

2 Beispielsweise wurde im Studiengang Betriebswirtschaftslehre der Wirtschaftsuniversität Prag die Bachelorarbeit abgeschafft und soll durch eine plagiatsicherere und alltagsnähere Prüfungsform ersetzt werden (<https://www.heise.de/news/ChatGPT-Co-Uni-schafft-Bachelorarbeiten-ab-9546851.html>, zuletzt abgerufen am 20.02.2024)

Domänen wie Programmierung ist eine tiefergehende Analyse von Lernprozessen noch aufwendig, unter anderem weil eine grosse Menge qualitativ hochwertiger Daten für das Training der Modelle benötigt wird (Cochran et al. 2022; Combéfis 2022). Es muss sich zukünftig zeigen, inwiefern allgemeine generative Sprachmodelle hier eine Abkürzung schaffen und für die Analyse von Lernprozessen angepasst werden können.

Dass generative KI zur Neubewertung bisheriger Prüfungsformate auffordert, kann insofern als positiv angesehen werden, als dass es bisherigen Reformversuchen Rückenwind verleihen könnte. Generative Sprachmodelle wirken hier als Verstärker bisheriger Diskurse, ohne gänzlich neue Probleme und Fragestellungen zu eröffnen. Dabei wird die Frage nach Prüfungsformen, die sowohl kompetenzorientiert als auch zuverlässig und mit begrenzten Ressourcen durchführbar sind, durch den Einzug von KI noch einmal betont.

Falls sich in einer Fachdisziplin die Kompetenzziele verändern oder auch überfachliche Kompetenzen stärker betont werden, sollte sich dies in den Prüfungsformen widerspiegeln. Prüfungsformen, die diese Kompetenzen alltagsnah überprüfen, gehen aktuell noch mit einem erhöhten Korrekturaufwand einher, welcher möglicherweise durch (teil-)automatisierte KI-Bewertungen erleichtert werden könnte. Formatives Assessment, welches darüber hinaus wünschenswert wäre, ist ebenfalls aus diesem Grund nicht leistbar. Die Fragen, wie formatives und möglicherweise auch summatives Assessment durch KI unterstützt werden kann, stellen sich weiterhin. Insofern verstärkt generative KI im Bereich der Prüfungen die bestehenden Fragen.

4. Lehr- und Lernmethode

Wenn also KI im Allgemeinen und generative Sprachmodelle im Speziellen eine grössere Bedeutung bei Kompetenzen und Prüfungen sowie Feedback erhalten und erhalten sollen, muss ihre Rolle auch in Lehr- und Lernsettings reflektiert werden. Wenn beispielsweise beim Verfassen wissenschaftlicher Texte zunehmend die KI-gestützte Literaturrecherche eine Rolle spielt (eine Veränderung in der Kompetenz), dann sollte die Fähigkeit, KI-gestützte Recherchewerkzeuge angemessen einzusetzen, nicht nur in der Prüfungssituation eingesetzt werden, sondern es sollte auch Gelegenheit zu Übung und Reflexion hierzu geben. Hierfür braucht es kostenlosen Zugang zu KI-Werkzeugen, um allen Studierenden und Lehrenden gleichermassen die Nutzung zu ermöglichen. So hat beispielsweise die University of Michigan verschiedene KI-Werkzeuge für alle Mitglieder der Universität freigeschaltet (Burns 2023).

Doch KI-Werkzeuge sollten nicht nur dort eingesetzt werden, wo sie unmittelbar zum Erreichen eines Lehr- und Lernziels mit KI-Bezug notwendig sind. Sie bieten darüber hinaus langfristig das Potenzial, Lehr- und Lernprozesse auf Arten und Weisen zu unterstützen, die zuvor aufgrund von begrenzten menschlichen Ressourcen nicht möglich waren.

Dieses Versprechen von KI-gestütztem Lernen gibt es indes schon lange und es wartet noch auf Einlösung (Bond et al. 2024; Zawacki-Richter et al. 2019). Vor dem Aufkommen allgemein verfügbarer generativer KI gab es zumindest im deutschsprachigen Raum keinen breiten Einsatz von KI in Lehr-Lern-Settings. Einzelne Anwendungen wie *Bettermarks*³ verwenden maschinelles Lernen für Adaptivität. Uns ist jedoch kein System bekannt, welches flächendeckend in Schulen oder Hochschulen eingesetzt wird. KI kommt teilweise in Intelligenten Tutoriellen Systemen, Learning Analytics oder unterstützenden Chatbots zum Einsatz (Bond et al. 2024; Mousavinasab et al. 2021). Oft sind diese Systeme jedoch noch regelbasiert, da wie beschrieben das Training von spezialisierten Modellen aufwendig ist. Insbesondere in Intelligenten Tutoriellen Systemen besteht die Aufgabe der KI auch nicht im «Schreiben», sondern im «Lesen», also in der automatischen Auswertung des Lernstands. Die Idee ist, dass Lernenden aufgrund dieser Analyse möglichst passende personalisierte Inhalte zugeteilt werden, wobei die Inhalte selbst bisher in der Regel von Menschen erstellt wurden.

Im Zuge von generativen Sprachmodellen wird nun häufig von *personalisiertem Lernen* gesprochen, welches die KI ermöglichen, obwohl generative KI den Lernstand höchstens rudimentär analysieren kann. Generative Sprachmodelle entwickeln kein *Lernermodell*, analysieren also nicht den aktuellen Wissensstand der Lernenden und geben nicht proaktiv die passenden Lerninhalte aus. Generative Sprachmodelle sind per se keine adaptiven Lernmedien. Das heisst, es werden keine Daten über die Lernenden für ein Lernermodell gesammelt und auch keine Entscheidungen für sie vom System getroffen.

Man kann generative KI also nicht als adaptives, wohl aber als adaptierbares Lernmedium verstehen. Die primäre Stärke von generativen Sprachmodellen liegt darin, dass Texte von Lernenden hinsichtlich Parametern wie Form, Länge, Sprache, Schwierigkeit oder inhaltlicher Fokus adaptiert werden können. Dies war zuvor mit keinem Werkzeug möglich. Zuvor mussten Lernende nach der für sie passenden Darstellungsform eines Inhalts suchen. Nun können sie beispielsweise ein Skript zusammenfassen lassen oder eine schwierige Textpassage vereinfacht erklären lassen.

Insofern erschliessen LLMs eine neue *Lernmethode*, also eine Art und Weise, sich gezielt Wissen und Fähigkeiten anzueignen: Lernende sind nicht mehr auf das Vorhandensein von Informationen in einer bestimmten Darstellungsform angewiesen,

3 <https://de.bettermarks.com>.

sondern können diese für die eigenen Bedürfnisse adaptieren. Vermutlich brauchen Lernende ein gewisses Mass an Unterstützung, um KI effektiv für diese Zwecke einzusetzen und die Limitationen zu kennen. Andernfalls besteht die Gefahr, dass sie mit zu wenig Vorwissen diese Werkzeuge nicht gezielt für das Lernen nutzen und nicht profitieren (Ben Youssef, Dahmani, und Ragni 2022). Dies ist auch problematisch, sofern KI insbesondere Lernende mit wenig Vorwissen unterstützen könnte, gleichzeitig aber auch Aufklärung über die Möglichkeiten und Limitationen von LLMs nötig ist. Nach einer Studie der *Boston Consulting Group* (Candelon et al. 2023) profitierten die Berater:innen mit weniger Vorwissen am stärksten in kreativen Aufgaben von der Arbeit mit ChatGPT. Sie verliessen sich jedoch auch stärker auf die Resultate und schnitten somit schlechter ab bei einer Aufgabe, bei der ChatGPT irreführende Resultate lieferte.

Bisher haben generative Sprachmodelle nicht zu gänzlich neuen *Lehrmethoden* geführt, sondern werden in bestehende Lehrmethoden integriert (Kasneci et al. 2023). Beispielsweise kann ein Chatbot eine Rolle einnehmen, die sonst leer bliebe, als Diskutant, Feedbackgeber, oder auch eine Rolle in einem Rollenspiel wie einem simulierten Beratungsgespräch. Auch Rollen, die zwischenmenschliche Kommunikation erleichtern, sind denkbar: KI könnte eine Diskussion am Ende eines Seminars zusammenfassen oder diese sogar moderieren. Letzteres scheint aktuell technisch noch schwierig, könnte mit weiteren multimodalen Fähigkeiten jedoch machbar werden.

Chatbots können also als weitere Akteure im Lernszenario eingesetzt werden, wobei diese Akteurhaftigkeit über das Einsetzen eines neuen Mediums hinausgeht. Mit Akteurhaftigkeit ist hierbei gemeint, dass generative Chatbots in einem grösseren Mass als bisherige Technologien dazu in der Lage sind, auf Eingaben der Nutzenden zu reagieren und eine zwischenmenschliche Interaktion zu imitieren. Generative Chatbots arbeiten nicht regelbasiert, das heisst, sie agieren nicht deterministisch, wodurch eine komplexe Interaktionssituation entsteht, die auch mittels Expertenwissen keiner eindeutig erkennbaren Ursache-Wirkung-Beziehung unterliegt. Auswirkungen der eigenen Inputs können lediglich im Nachhinein teilweise rekonstruiert werden (Moebert 2022). Darüber hinaus kommen laufend neue Modelle auf den Markt, die sich jeweils anders verhalten, sodass neu ausgetestet werden muss, welche Veränderungen in der Interaktion sich ergeben. Im Hinblick auf die Akteurhaftigkeit stellt sich die Frage, welches Mass an Anthropomorphisierung wünschenswert ist (Bender 2022), zumal ein humanoides Chatbot-Interface darüber hinwegtäuschen kann, dass potenziell alle Daten der Interaktion an die dahinterstehenden Firmen gehen. Hier besteht die Aufgabe, (offene) Sprachmodelle zu fördern und bereitzustellen, die idealerweise lokal auf Endgeräten laufen, sodass die Nutzenden die Kontrolle über Datenflüsse und Schnittstellen haben.

Eine weitere Fragestellung ist, welche sozialen Implikationen die KI als zusätzlicher Akteur verursacht. Wie verändern sich Interaktionen zwischen Lehrenden und Lernenden in einem Dreieck mit KI? Inwiefern verschiebt sich die Verantwortung für den Lernprozess?

Sich in Lehr-Lern-Settings mit den Affordanzen dieses neuen Werkzeugs beziehungsweise des neuen Akteurs auseinanderzusetzen, ist gemäss dem *Constructive Alignment* ein notwendiger Baustein angesichts der Veränderungen im Bereich der Kompetenzen und Prüfungen. Indem Lernende und Lehrende generative KI in verschiedenen Kontexten ausprobieren und darüber reflektieren, werden Kompetenzen in diesem neuen Lernfeld aufgebaut. In Prüfungen können Lernende zeigen, wie sie generative KI einsetzen oder darauf verzichten. Aber KI kann nicht nur als Werkzeug und damit als Lerninhalt eine Rolle spielen, sondern auch an Stellen unterstützen, die bisher aufgrund begrenzter Ressourcen unausgefüllt blieben. In Lehr-Lern-Settings wird mit zunehmender Multimodalität die Akteurhaftigkeit von generativen KI-Systemen deutlich und es stellen sich Fragen zu den Eigenschaften dieser neuen Akteure und zu den Möglichkeiten und Grenzen ihres Einsatzes.

5. Zusammenfassung und Diskussion

Auch im Hinblick auf generative KI zeigt sich entsprechend dem Constructive Alignment die didaktische Verschränkung von Kompetenzziele, Prüfungsformaten und didaktischen Entscheidungen. Wenn KI-relevante Fähigkeiten, Wissen und Einstellungen in die Curricula integriert werden sollen, zieht dies auch notwendigerweise eine Integration in Prüfungsszenarien und Lehr-Lern-Settings nach sich. Neben der Rolle von KI als Lerngegenstand gibt es darüber hinaus in diesen Bereichen Fragestellungen zum Einsatz von KI als Werkzeug, Medium oder Akteur.

Und doch unterscheiden sich die Bereiche dahingehend, inwiefern sich durch generative KI überhaupt neue Fragen stellen. Im Bereich der Kompetenzen stellen sich alte und neue Fragen. Die Fragen nach *AI Literacy* schliessen an Diskussionen über «digitale Kompetenz» und «Future Skills» an. Als neue Frage kommt das Phänomen des *Deskilling*s hinzu, welches in diesem Mass noch nicht in seiner Breite diskutiert wurde. Im Bereich der Prüfungen scheinen sich in erster Linie alte Fragen mit neuer Dringlichkeit zu stellen: Wie sollten Prüfungsformate gestaltet sein, damit sie tatsächlich die Kompetenzzuwächse erheben, die als Lehr- und Lernziele gedacht waren? Wie sehen alternative Prüfungsformate aus, die mit begrenzten Ressourcen umsetzbar und valide sind? In Bezug auf die didaktischen Lehr-Lernsettings kommt es durch LLMs nicht zu gänzlich neuen Lern- und Lehrmethoden. Gleichzeitig bieten LLMs zwei Potenziale, die es zu untersuchen gilt: Zum einen bieten sie Lernenden die Möglichkeit, jegliche Lernmedien für die eigenen Zwecke anzupassen oder komplett zu generieren. Damit wird das Lernangebot, welches durch die Lehrenden gemacht

wird, stärker durch Lernende adaptierbar als zuvor. Zum anderen können generative Sprachmodelle als weiterer Akteur in Lernsettings an Stellen eingesetzt werden, die zuvor nicht oder von Menschen besetzt waren, z. B. in Dialogen, Peer-Teaching, Moderation oder Feedback. Dadurch kann sich nicht nur verändern, wie oft diese didaktischen Settings umgesetzt werden, sondern sie verändern auch ihre Gestalt. Forschung zu Conversational Agents und Pedagogical Agents kann hier Antworten auf Fragen bieten, wie Lernszenarien mit diesen neu verfügbaren Akteuren ausgestaltet werden sollten (Baker 2016; Richards und Dignum 2019).

In diesen teils neuen, teils alten, aber verstärkten Fragestellungen zeigt sich, wie KI als Verstärker auf ein System wirken kann (Muuß-Merholz 2019). Angesichts dieser Verstärkerwirkung müssen wir uns noch dringender fragen, auf welche Ziele die Bildungssysteme ausgerichtet sind. Ein Beispiel für eine Zielvariable von Hochschulbildung ist beruflicher Erfolg (vgl. Pellegrino und Hilton 2012). Ein Fokus auf dieses Ziel hat Auswirkungen auf Kompetenzziele: Sie sollen beispielsweise eine zukunftssichere Employability ermöglichen. Prüfungsformate sollen sich dementsprechend stärker an beruflichen Kontexten ausrichten und die Lehrmethoden könnten demgemäß stärker projektbasiert orientiert sein. Wenn diese Zielvariable nicht explizit diskutiert, sondern implizit angenommen wird, verstärkt KI diese Entwicklungen, Bildung im Hinblick auf arbeitsmarktorientierte Verwertbarkeit und Selbstoptimierung auszurichten (Kalz 2023; Kerres 2023).

Auch übergreifende ethische, politische, soziale und rechtliche Fragen werden verstärkt (Leschke und Salden 2023). Wie sieht Fairness in diesen Systemen aus? Wie können KI-Systeme zur Bildungsgerechtigkeit beitragen oder Ungerechtigkeit verschärfen? Wie können rechtliche Rahmenbedingungen dazu geschaffen werden? Welche Gefahren bestehen, wenn wir das Thema ignorieren? Wie stark sollen kommerzielle Partner im Bildungsraum wirken? Welche Rechte haben Lernende in Bezug auf ihre Daten und Erzeugnisse? Sollten wir im Sinne des gesellschaftlichen Bildungsauftrags komplett auf offene Modelle setzen? Kann KI als Technologie der Anlass oder auch der Hebel sein, über neue Lernkulturen nachzudenken?

Der Diskurs zeigt auch, dass noch viele Fragen ungeklärt sind. Dem kann man auf zwei Ebenen begegnen: Zum einen sollte mehr Raum zum Experimentieren und kritischen Reflektieren an Hochschulen geschaffen werden und Erkenntnisse daraus sollten iterativ in die Community zurückgespielt werden. Gerade in Anbetracht der Geschwindigkeit der Entwicklung ist ein niedrigschwelliges, gemeinsames Entwickeln und Gestalten zwischen Studierenden, Lehrenden und anderen Stakeholdern essenziell. Als Voraussetzung und Rahmung sollten dafür aber auch Zugänge zu KI-Modellen, Infrastruktur und Weiterbildung geschaffen werden (Schleiss et al. 2023).

Zum anderen ist auch die Forschung gefragt, vor allem in Bezug auf Lernwirksamkeit, die Gestaltung neuer Lernszenarien und die kritische Begleitung (Bond et al. 2024). Dies sollte mit einem stärkeren Theoriebezug geschehen als bisher

(Zawacki-Richter et al. 2019). Abschliessend lässt sich festhalten: KI ist nicht als deterministischer Faktor zu sehen, der von sich aus auf die Bildung einwirkt. Stattdessen gilt es, auf allen Ebenen der Forschung und der Praxis reflektiert und kritisch, aber auch offen und experimentierfreudig die neuen Möglichkeiten durch generative KI zu prüfen und zu erproben.

Literatur

- Baker, Ryan S. 2016. «Stupid Tutoring Systems, Intelligent Humans». *International Journal of Artificial Intelligence in Education* 26 (2): 600–14. <https://doi.org/10.1007/s40593-016-0105-0>.
- Ben Youssef, Adel, Mounir Dahmani, und Ludovic Ragni. 2022. «ICT Use, Digital Skills and Students' Academic Performance: Exploring the Digital Divide». *Information* 13 (3): 129. <https://doi.org/10.3390/info13030129>.
- Bender, Emily M. 2022. «Human-like Programs Abuse Our Empathy – Even Google Engineers Aren't Immune». *The Guardian*, 14. Juni 2022, Abschn. Opinion. <https://www.theguardian.com/commentisfree/2022/jun/14/human-like-programs-abuse-our-empathy-even-google-engineers-arent-immune>.
- Benotti, Luciana, Mara Cecilia Martnez, und Fernando Schapachnik. 2018. «A Tool for Introducing Computer Science with Automatic Formative Assessment». *IEEE Transactions on Learning Technologies* 11 (2): 179–92. <https://doi.org/10.1109/TLT.2017.2682084>.
- Biggs, John. 1996. «Enhancing Teaching through Constructive Alignment». *Higher Education* 32 (3): 347–64. <https://doi.org/10.1007/BF00138871>.
- Bond, Melissa, Hassan Khosravi, Maarten De Laat, Nina Bergdahl, Violeta Negrea, Emily Oxley, Phuong Pham, Sin Wang Chong, und George Siemens. 2024. «A Meta Systematic Review of Artificial Intelligence in Higher Education: A Call for Increased Ethics, Collaboration, and Rigour». *International Journal of Educational Technology in Higher Education* 21 (1): 4. <https://doi.org/10.13140/RG.2.2.31921.56162/1>.
- Bozkurt, Aras, und Ramesh C. Sharma. 2023. «Generative AI and Prompt Engineering: The Art of Whispering to Let the Genie Out of the Algorithmic World». <https://doi.org/10.5281/ZENODO.8174941>.
- Brightwell, Richard, Janine-Helen Daniel, und Angus Stewart. 2004. «Evaluation: is an open book examination easier?» *Bioscience Education* 3 (1): 1–10. <https://doi.org/10.3108/beej.2004.03000004>.
- Buck, Isabella, und Anika Limburg. 2023. «Hochschulbildung vor dem Hintergrund von Natural Language Processing (KI-Schreibtools). Ein Framework für eine zukunftsfähige Lehr- und Prüfungspraxis» 9 (1): 70. <https://doi.org/10.3278/HSL2306W>.
- Budde, Jannica, Jens Tobor, und Jasper Beyermann. 2023. «Blickpunkt Digitale Prüfungen». Hochschulforum Digitalisierung.

- Burns, Tom. 2023. «ITS debuts custom artificial intelligence services across U-M». *The University Record* (blog). 21. August 2023. <https://record.umich.edu/articles/its-debuts-customized-ai-services-to-u-m-community/>.
- Candelon, Francois, Lisa Krayner, Saran Rajendran, und David Zuluaga Martinez. 2023. «How People Can Create – and Destroy – Value with Generative AI». BCG Global. 21. September 2023. <https://www.bcg.com/publications/2023/how-people-create-and-destroy-value-with-gen-ai>.
- Chase, Kiera, und Dor Abrahamson. 2015. «Reverse-Scaffolding Algebra: Empirical Evaluation of Design Architecture». *ZDM* 47 (7): 1195–1209. <https://doi.org/10.1007/s11858-015-0710-7>.
- Cochran, Keith, Clayton Cohn, Nicole Hutchins, Gautam Biswas, und Peter Hastings. 2022. «Improving Automated Evaluation of Formative Assessments with Text Data Augmentation». In *Artificial Intelligence in Education*, herausgegeben von Maria Mercedes Rodrigo, Noburu Matsuda, Alexandra I. Cristea, und Vania Dimitrova, 390–401. Lecture Notes in Computer Science. Cham: Springer. https://doi.org/10.1007/978-3-031-11644-5_32.
- Combéfis, Sébastien. 2022. «Automated Code Assessment for Education: Review, Classification and Perspectives on Techniques and Tools». *Software* 1 (1): 3–30. <https://doi.org/10.3390/software1010002>.
- Dunning, David. 2011. «The Dunning–Kruger Effect: On Being Ignorant of One’s Own Ignorance». In *Advances in Experimental Social Psychology*, herausgegeben von James M. Olson und Mark P. Zanna, 44: 247–96. Academic Press. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>.
- Ebbinghaus, Uwe. 2023. «KI in der Schule: Enkelejda Kasneci über Potential und Risiken». *FAZ.NET*, 31. Mai 2023. <https://www.faz.net/aktuell/karriere-hochschule/ki-in-der-schule-enkelejda-kasneci-ueber-potential-und-risiken-18911332.html>.
- Ehlers, Ulf-Daniel. 2020. «Future Skills für die Welt von morgen». In *Future Skills*, von Ulf-Daniel Ehlers, 57–100. *Zukunft der Hochschulbildung – Future Higher Education*. Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-29297-3_5.
- Engel, Juliane, und Michael Kerres. 2023. «Bildung in der Nächsten Gesellschaft – Eine post-digitale Sicht auf neue Formen der Subjektivierung». *Ludwigsburger Beiträge zur Medienpädagogik* 23 (Oktober): 1–13. <https://doi.org/10.21240/lbzm/23/04>.
- Gottschling, Steffen, Tobias Seidl, und Cornelia Vonhof. 2023. «Nutzung von KI-Tools durch Studierende. Eine exemplarische Untersuchung studentischer Nutzungsszenarien». (Under Review) *Die Hochschullehre*. <https://hochschulforumdigitalisierung.de/blog/hey-chatgpt/>.
- Grinschgl, Sandra, Frank Papenmeier, und Hauke S Meyerhoff. 2021. «Consequences of Cognitive Offloading: Boosting Performance but Diminishing Memory». *Quarterly Journal of Experimental Psychology* 74 (9): 1477–96. <https://doi.org/10.1177/17470218211008060>.
- Hamilton, Arran, Dylan Wiliam, und John Hattie. 2023. «The Future of AI in Education: 13 Things We Can Do to Minimize the Damage». <https://doi.org/10.35542/osf.io/372vr>.

- He, Gaole, Lucie Kuiper, und Ujwal Gadiraju. 2023. «Knowing About Knowing: An Illusion of Human Competence Can Hinder Appropriate Reliance on AI Systems». In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. CHI '23. New York, NY, USA: Association for Computing Machinery. <https://doi.org/10.1145/3544548.3581025>.
- Hu, Krystal, und Krystal Hu. 2023. «ChatGPT Sets Record for Fastest-Growing User Base – Analyst Note». *Reuters*, 2. Februar 2023, Abschn. Technology. <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>.
- Ifenthaler, Dirk, und Hendrik Drachler. 2020. «Learning Analytics Spezielle Forschungsmethoden in der Bildungstechnologie». In *Handbuch Bildungstechnologie: Konzeption und Einsatz digitaler Lernumgebungen*, herausgegeben von Helmut Niegemann und Armin Weinberger. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-662-54368-9>.
- Jumper, John, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, et al. 2021. «Highly Accurate Protein Structure Prediction with AlphaFold». *Nature* 596 (7873): 583–89. <https://doi.org/10.1038/s41586-021-03819-2>.
- Kalz, Marco. 2023. «Zurück in die Zukunft? Eine literaturbasierte Kritik der Zukunftskompetenzen». *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* (Occasional Papers): 332–52. <https://doi.org/10.21240/mpaed/00/2023.11.19.X>.
- Kasneci, Enkelejda, Kathrin Sessler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, et al. 2023. «ChatGPT for good? On opportunities and challenges of large language models for education». *Learning and Individual Differences* 103 (April): 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Kerres, Michael. 2017. «Digitalisierung als Herausforderung für die Medienpädagogik: Bildung in einer digital geprägten Welt». In *Pädagogischer Mehrwert? Digitale Medien in Schule und Unterricht*, herausgegeben von Christian Fischer. Münster: Waxmann.
- Kerres, Michael. 2023. «Bildung in der digitalen Welt: (Wie) Kann digitale Kompetenz vermittelt werden?» In *Lernen in und für die Schule im Zeitalter der Digitalität – Chancen und Herausforderungen für Lehrkräfte, Unterricht und Lernende*, herausgegeben von N. McElvany, R. Lorenz, M. Becker, H. Gaspard, und F. Laueremann. Bd. 6. IFS-Bildungsdialoge. Münster: Waxmann.
- Khalil, Mohammad, und Erkan Er. 2023. «Will ChatGPT get you caught? Rethinking of Plagiarism Detection». arXiv. <https://doi.org/10.48550/arXiv.2302.04335>.
- Leber, Jens, Anna Heinemann, Pia Sander, Chirine Ahmad, Lara Meneghinello, und Louis Peters. 2023. «Informationskompetenz von Schüler:innen stärken: Wie kann man das Prüfen von Informationen durch den Einsatz digitaler Medien fördern?» In *Bildung im digitalen Wandel. Die Bedeutung digitaler Medien für soziales Lernen und Teilhabe*, herausgegeben von Michael Kerres, Sybille Stöbe-Blossey, Ulrike Cress, Marc Rittberger, und Josef Schrader, 4:75–106. Digitalisierung in der Bildung. <https://www.waxmann.com/index.php?eID=download&buchnr=4846>.
- Leschke, Jonas, und Peter Salden. 2023. *Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung*. Ruhr-Universität Bochum.

- Liang, Weixin, Mert Yuksekogonul, Yining Mao, Eric Wu, und James Zou. 2023. «GPT Detectors Are Biased against Non-Native English Writers». *Patterns* 4 (7). <https://doi.org/10.1016/j.patter.2023.100779>.
- Long, Duri, und Brian Magerko. 2020. «What Is AI Literacy? Competencies and Design Considerations». In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. Honolulu HI USA: ACM. <https://doi.org/10.1145/3313831.3376727>.
- Malinka, Kamil, Martin Peresíni, Anton Firc, Ondrej Hujnák, und Filip Janus. 2023. «On the Educational Impact of ChatGPT: Is Artificial Intelligence Ready to Obtain a University Degree?» In *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education V. 1*, 47–53. Turku Finland: ACM. <https://doi.org/10.1145/3587102.3588827>.
- Moebert, Tobias. 2022. *Zum Einfluss von Adaptivität auf die Interaktion mit Bildungstechnologien*. Gesellschaft für Informatik e.V. <https://dl.gi.de/handle/20.500.12116/38841>.
- Monaghan, John. 2016. «The Calculator Debate». In *Tools and Mathematics*, herausgegeben von John Monaghan, Luc Trouche, und Jonathan M. Borwein, 305–31. Mathematics Education Library. Cham: Springer. https://doi.org/10.1007/978-3-319-02396-0_13.
- Mousavinasab, Elham, Nahid Zarifsanaiey, Sharareh R. Niakan Kalhori, Mahnaz Rakhshan, Leila Keikha, und Marjan Ghazi Saeedi. 2021. «Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods». *Interactive Learning Environments* 29 (1): 142–63. <https://doi.org/10.1080/10494820.2018.1558257>.
- Muuß-Merholz, Jöran. 2019. «Der große Verstärker. Spaltet die Digitalisierung die Bildungswelt? – Essay». J&K – Jöran und Konsorten. 29. November 2019. <https://www.joeran.de/der-grosse-verstaerker-spaltet-die-digitalisierung-die-bildungswelt-essay/>.
- Paiva, José Carlos, José Paulo Leal, und Álvaro Figueira. 2022. «Automated Assessment in Computer Science Education: A State-of-the-Art Review». *ACM Transactions on Computing Education* 22 (3): 34 (1):34–40. <https://doi.org/10.1145/3513140>.
- Pellegrino, James W., und Margaret L. Hilton. 2012. *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.
- Rafner, Janet, Dominik Dellermann, Arthur Hjorth, Dóra Verasztó, Constance Kampf, Wendy Mackay, und Jacob Sherson. 2021. «Deskilling, Upskilling, and Reskilling: A Case for Hybrid Intelligence». *Morals & Machines* 1 (2): 24–39. <https://doi.org/10.5771/2747-5174-2021-2-24>.
- Ramesh, Dadi, und Suresh Kumar Sanampudi. 2022. «An Automated Essay Scoring Systems: A Systematic Literature Review». *Artificial Intelligence Review* 55 (3): 2495–2527. <https://doi.org/10.1007/s10462-021-10068-2>.
- Reinmann, Gabi. 2023. «Deskilling durch Künstliche Intelligenz?» Diskussionspapier 25. Hochschulforum Digitalisierung. https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_DP_25_Deskilling.pdf.
- Richards, Deborah, und Virginia Dignum. 2019. «Supporting and Challenging Learners through Pedagogical Agents: Addressing Ethical Issues through Designing for Values». *British Journal of Educational Technology* 50 (6): 2885–2901. <https://doi.org/10.1111/bjet.12863>.

- Schleiss, Johannes, Matthias Carl Laupichler, Tobias Raupach, und Sebastian Stober. 2023. «AI Course Design Planning Framework: Developing Domain-Specific AI Education Courses». *Education Sciences* 13 (9): 954. <https://doi.org/10.3390/educsci13090954>.
- Schleiss, Johannes, Dana-Kristin Mah, Katrin Böhme, David Fischer, Janne Mesenhöller, Benjamin Paaßen, Sabrina Schork, und Johannes Schrumpf. 2023. «Künstliche Intelligenz in der Bildung – Drei Zukunftsszenarien und fünf Handlungsfelder. KI-Campus». *KI Campus*. <https://osnascholar.uni-osnabrueck.de/handle/unios/70865>.
- Schüller, Katharina, Florian Rampelt, Henning Koch, und Johannes Schleiss. 2023. «Better Ready than Just Aware: Data and AI Literacy as an Enabler for Informed Decision Making in the Data Age». *INFORMATIK 2023 - Designing Futures: Zukünfte gestalten* 425-430. https://doi.org/10.18420/inf2023_49.
- Stanja, Judith, Wolfgang Gritz, Johannes Krugel, Anett Hoppe, und Sarah Dannemann. 2023. «Formative Assessment Strategies for Students' Conceptions – The Potential of Learning Analytics». *British Journal of Educational Technology* 54 (1): 58–75. <https://doi.org/10.1111/bjet.13288>.
- Von Garrel, Jörg, und Jana Mayer. 2023. «Artificial Intelligence in Studies – Use of ChatGPT and AI-Based Tools among Students in Germany». *Humanities and Social Sciences Communications* 10 (1): 799. <https://doi.org/10.1057/s41599-023-02304-7>.
- Vuorikari, Riina, Stefano Kluzer, und Yves Punie. 2022. *DigComp 2.2, The Digital Competence Framework for Citizens: With New Examples of Knowledge, Skills and Attitudes*. LU: Publications Office of the European Union. <https://data.europa.eu/doi/10.2760/115376>.
- Weinert, Franz E. 2001. «Competencies and Key Competencies: Educational Perspective». In *International Encyclopedia of the Social & Behavioral Sciences*, herausgegeben von Neil J. Smelser und Paul B. Baltes, 2433–36. Oxford: Pergamon. <https://doi.org/10.1016/B0-08-043076-7/02384-6>.
- Weizenbaum, Joseph. 1976. *Computer power and human reason: From judgment to calculation*. Computer power and human reason: From judgment to calculation. Oxford, England: W. H. Freeman.
- Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond, und Franziska Gouverneur. 2019. «Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators?» *International Journal of Educational Technology in Higher Education* 16 (1): 39. <https://doi.org/10.1186/s41239-019-0171-0>.
- Zhai, Xiaoming, Peng He, und Joseph Krajcik. 2022. «Applying Machine Learning to Automatically Assess Scientific Models». *Journal of Research in Science Teaching* 59 (10): 1765–94. <https://doi.org/10.1002/tea.21773>.