

phpublico

Fachzeitschrift für Bildung und Erziehung

Empfehlungen für eine evidenzbasierte Nutzung von Künstlicher Intelligenz und Large Language Modells in Hochschulen: Ergebnisse eines systematischen Reviews

Thomas Leitgeb¹ & Michael Leitgeb¹

¹Private Pädagogische Hochschule Burgenland

ORCID Thomas Leitgeb: <https://orcid.org/0000-0002-0678-5756>

ORCID Michael Leitgeb: <https://orcid.org/0009-0001-8839-4926>

Korrespondierender Autor: Thomas Leitgeb (thomas.leitgeb@ph-burgenland.at)

Abstract Deutsch

Die Künstliche Intelligenz (KI) und insbesondere Large Language Models (LLMs) haben in den letzten Jahren erheblich an Bedeutung im Bildungssektor gewonnen. Diese Technologien bieten vielfältige Einsatzmöglichkeiten, wie die Personalisierung des Lernens, Effizienzsteigerungen in administrativen Prozessen und die Einführung innovativer Lehrmethoden. Trotz der zahlreichen Potenziale besteht eine Forschungslücke hinsichtlich der konkreten Gestaltung und Implementierung von KI-Systemen in der Hochschulbildung, die sowohl technische Effizienz als auch ethische Standards berücksichtigt. Die vorliegende Arbeit schließt diese Lücke durch eine systematische Literaturanalyse gemäß den PRISMA-Richtlinien, wobei 116 relevante Studien aus den Datenbanken Web of Science, Scopus und Google Scholar ausgewertet wurden.

Die Analyse zeigt, dass KI und LLMs signifikante Vorteile für die Hochschulbildung bieten, wie die Erstellung individualisierter Lernpfade und die Automatisierung von Bewertungsprozessen. Gleichzeitig identifiziert die Studie wesentliche Herausforderungen, darunter Datenschutzprobleme, algorithmische Verzerrungen (Bias) und die mangelnde Erklärbarkeit komplexer Modelle. Ethische Implikationen wie Fairness und Gleichberechtigung werden ebenfalls kritisch beleuchtet. Um diese Herausforderungen zu bewältigen, werden konkrete Handlungsempfehlungen für hochschulische Bildungseinrichtungen entwickelt. Dazu gehören die Entwicklung hochschuleigener KI-Systeme zur besseren Datenkontrolle, die Förderung interdisziplinärer Zusammenarbeit, umfassende Schulungsprogramme für Lehrkräfte sowie die Etablierung von Governance-Strukturen und ethischen Leitlinien.

Die Ergebnisse unterstreichen die Notwendigkeit einer integrativen Perspektive, die technische und ethische Aspekte gleichermaßen berücksichtigt, um das volle Potenzial von KI im Bildungsbereich verantwortungsvoll auszuschöpfen. Diese Arbeit liefert praxisorientierte Leitlinien für die effektive und ethisch fundierte Implementierung von KI-Technologien in der Hochschulbildung und trägt somit zur Verbesserung der Bildungsqualität und zur Sicherstellung von Datenschutz und Fairness bei.

Abstract English

Artificial Intelligence (AI), particularly Large Language Models (LLMs), have gained significant importance in the education sector in recent years. These technologies offer a variety of applications, such as the personalization of learning, efficiency improvements in administrative processes, and the introduction of innovative teaching methods. Despite the numerous potentials, there exists a research gap regarding the concrete design and implementation of AI systems in higher education that consider both technical efficiency and ethical standards. The present work closes this gap through a systematic literature analysis in accordance with PRISMA guidelines, evaluating 116 relevant studies from the Web of Science, Scopus, and Google Scholar databases.

The analysis shows that AI and LLMs provide significant advantages for higher education, such as the creation of individualized learning paths and the automation of assessment processes.

At the same time, the study identifies essential challenges, including data privacy issues, algorithmic biases, and the lack of explainability of complex models. Ethical implications, such as fairness and equality, are also critically examined. To address these challenges, concrete recommendations for educational institutions are developed. These include the development of proprietary, internally hosted AI systems for better data control, the promotion of interdisciplinary collaboration, comprehensive training programs for educators, as well as the establishment of governance structures and ethical guidelines.

The results emphasize the necessity of an integrative perspective that equally considers technical and ethical aspects to responsibly harness the full potential of AI in the educational sector. This work provides practical guidelines for the effective and ethically grounded implementation of AI technologies in higher education, thereby contributing to the improvement of educational quality and ensuring data privacy and fairness.

Einleitung

Künstliche Intelligenz (KI) hat in den letzten Jahren erheblich an Bedeutung gewonnen und beeinflusst verschiedene gesellschaftliche Bereiche, insbesondere den Bildungssektor (Kasneji et al., 2023). Künstliche Intelligenz wird definiert als die Fähigkeit einer Maschine, menschliche Fähigkeiten wie logisches Denken, Lernen, Planen und Kreativität zu imitieren (Europäisches Parlament, 2023). Während KI ursprünglich vor allem in technischen Disziplinen entwickelt wurde, ist sie mittlerweile zu einer Schlüsseltechnologie geworden, die tiefgreifende Veränderungen in zahlreichen wissenschaftlichen Feldern ermöglicht (Russel & Norvig, 2021). Im Bildungsbereich bietet KI somit ebenfalls vielseitige Einsatzmöglichkeiten (z. B. Zawacki-Richter et al., 2019). Besonders hervorzuheben ist dabei die Nutzung generativer KI-Systeme wie Large Language Models (LLMs), die durch ihre Fähigkeit, menschliche Sprache zu verstehen und zu erzeugen, neue Anwendungsmöglichkeiten eröffnen (Singh et al., 2024). Trotz der zahlreichen Potenziale von KI und LLMs im Bildungsbereich, die bereits in zahlreichen Studien hervorgehoben wurden (z. B. Chen et al., 2020; Daun et al., 2024), ist die rechtskonforme Nutzung von KI-Systemen und LLMs problematisch (Vincent-Lancrin & Vlies, 2020; Singhal, 2024). Vor allem im datensensiblen Bildungsbereich besteht eine Forschungslücke hinsichtlich der konkreten Gestaltung und Implementierung von auf LLMs basierenden KI-Systemen, die sowohl technische Effizienz als auch ethische Standards an Hochschulen berücksichtigen. Mit der vorliegenden Arbeit wird versucht, diese Forschungslücke anhand einer systematischen Literaturanalyse zu schließen und aus der Synthese Empfehlungen abzuleiten.

Theoretische Grundlagen und Forschungsstand

Bisherige Untersuchungen konzentrieren sich häufig entweder auf die technischen Vorteile von KI im Bildungswesen oder auf die ethischen Implikationen, ohne eine integrative Perspektive zu bieten (Holmes et al., 2019; Selwyn, 2019). Insbesondere fehlt es an detaillierten Analysen, wie die Herausforderungen der ‚Halluzinationen‘ von LLMs¹ (Valentin et al., 2024) und der mangelnden Erklärbarkeit komplexer Modelle (Lipton, 2018) überwunden werden können. Während einige Studien die Relevanz des Datenschutzes betonen (Drachsler & Greller, 2016; Slade & Prinsloo, 2013), bleibt unklar, wie Bildungseinrichtungen praktische Lösungen implementieren können, um entsprechende Bedenken effektiv anzugehen. Ebenso gibt es einen Mangel an konkreten Empfehlungen zur Vermeidung von Bias² in KI-Systemen, die die Chancengleichheit im Bildungswesen beeinträchtigen könnten (Omughelli et al., 2024).

Bereits einige systematische Literaturanalysen setzen sich mit dem Einsatz von KI bzw. LLMs an Hochschulen auseinander.

¹ Halluzinationen in großen Sprachmodellen (LLMs) bezeichnen die Generierung von inhaltlich falschen oder erfundenen Aussagen, die nicht auf den Trainingsdaten basieren, aber dennoch kohärent erscheinen. (Wang et al, 2024)

² Bias in der künstlichen Intelligenz bezieht sich auf systematische Verzerrungen in den Entscheidungen oder Ergebnissen von KI-Systemen, die durch unausgewogene oder nicht repräsentative Trainingsdaten entstehen. Diese Verzerrungen können dazu führen, dass bestimmte Gruppen ungerechtfertigt bevorzugt oder diskriminiert werden, basierend auf Merkmalen wie Geschlecht, Ethnie, sozialem Status oder anderen demografischen Eigenschaften (Bhattacharya, 2024).

Crompton und Burke (2023) fertigten eine systematisch Literaturübersicht zu KI-Anwendungen im Hochschulbereich an, basierend auf 138 Studien aus dem Zeitraum von 2016 bis 2022. Sie analysierten die geografische Verteilung, den Forschungskontext sowie die Art und den Umfang der eingesetzten KI mit besonderem Fokus auf die Auswirkungen von KI auf Lehren und Lernen. Die Autor:innen unterteilen die KI-Anwendungen in vier Hauptkategorien: Bewertung und Beurteilung (z. B. automatische Bewertung und automatisches Feedback), Vorhersage akademischer Leistungen, intelligente Assistenten wie Chatbots sowie Studierendenverwaltung durch Lernmustererkennung und Profilbildung. Diese Übersicht liefert wertvolle Einblicke in die vielfältigen Einsatzmöglichkeiten von KI in der Hochschulbildung und betont gleichzeitig, dass wenige Studien aus dem deutschsprachigen Raum vorliegen.

Zawacki-Richter et al. (2019) bieten einen systematischen Überblick zur Forschung über KI im Hochschulbereich auf Basis von 146 Arbeiten aus den Jahren 2007 bis 2018. Der Fokus lag auf Definitionen von KI, ethischen Implikationen sowie der Anwendung von KI in der Hochschulbildung. Nur 1,4 % der untersuchten Studien befassten sich demnach mit ethischen Fragen, insbesondere solchen zu Datenschutz, Kosten und dem zeitlichen Aufwand für die Implementierung von KI-basierten Methoden. Ein Großteil der Arbeiten (58) behandelte den Einsatz von KI zur Erstellung von Studierendenprofilen, zur Unterstützung bei Zulassungsentscheidungen sowie zur Vorhersage von Lernverläufen und von Studienabbrüchen. In weiteren 29 Studien wurde der Einsatz intelligenter Tutorensysteme analysiert, die Wissenslücken diagnostizieren und automatisiertes Feedback geben. 36 Arbeiten befassten sich mit automatisierten Bewertungsmethoden, darunter Benotung und Beurteilung von Engagement und akademischer Integrität. Zusätzlich thematisierten 27 Arbeiten die Nutzung von KI zur Personalisierung des Lernens und zur Unterstützung von Lehrkräften bei der Unterrichtsgestaltung. Diese Forschung verdeutlicht sowohl die Potenziale als auch die Herausforderungen, die mit KI im Hochschulbereich einhergehen, insbesondere in Bezug auf ethische Fragen und die Integration in Bildungsprozesse.

Chu et al. (2020) analysierten die 50 meistzitierten Artikel über KI im Hochschulbereich, die im *Web of Science* veröffentlicht worden waren. Dabei lag der Fokus auf der Art und dem Umfang der in den untersuchten Artikeln angewandten KI, was in direktem Zusammenhang mit dem Themenkomplex ‚Auswirkungen von KI auf das Lehren und Lernen im Hochschulbereich‘ steht. Sie betrachteten auch den Kontext, in dem die Studien durchgeführt wurden, sowie die Hauptthemen, insbesondere die verwendeten Algorithmen, und identifizierten die produktivsten Autor:innen. Der Untersuchungszeitraum umfasste die Jahre 1996 bis 2020. Obwohl die meisten Ergebnisse auf einen technischen Bereich verweisen, wurde KI in verschiedenen Kontexten eingesetzt, darunter Empfehlungssysteme (fünf Arbeiten), prädiktive Systeme, etwa zur Identifikation von Studierenden mit Studienabbruchsrissen (22 Arbeiten), und intelligente Tutorensysteme, die automatisch die Bedürfnisse der Studierenden diagnostizieren und personalisierte Interventionen anbieten (23 Arbeiten). Die Autor:innen betonen die Notwendigkeit einer stärker pädagogisch geleiteten Implementierung von KI und fordern mehr Fortbildungen für Lehrkräfte.

Eine systematische Literaturanalyse des Forums Neue Medien Austria (FNMA) – durchgeführt von Pishtari, Wagner und Ley (2024) – bietet einen umfassenden Forschungsüberblick zum Einsatz von KI in der Hochschulbildung. In dem Bericht werden vor allem die Nutzung von KI für das Lehren und Lernen beleuchtet und systematische Literaturübersichten sowie aktuelle empirische Studien aus dem Jahr 2023 untersucht. Zunächst werden systematische Übersichten

identifiziert, die auf die Auswirkungen von KI auf Lehr- und Lernpraktiken fokussieren, einschließlich der Akzeptanz von KI durch Lehrende und Lernende. Die Autor:innen untersuchen die Einsatzgebiete von KI beispielsweise als Vorhersagesysteme zur Erkennung gefährdeter Studierender, als intelligente Tutorensysteme sowie zur Personalisierung von Lernressourcen. Dabei zeigt sich, dass KI vor allem im Bereich der Lernanalyse, zur Unterstützung von Lernprozessen und bei der Erstellung von Feedback eingesetzt wird. Mit der Untersuchung wird hervor gehoben, dass der Einsatz von KI in der Hochschullehre weiterhin Herausforderungen wie ethische Bedenken und Fragen des Datenschutzes mit sich bringt. In dem Bericht wird zudem eine stärkere Integration von Bildungs- und Lerntheorien in die KI-Implementierung gefordert und betont, dass weitere empirische Forschung zur langfristigen Wirkung von KI-Anwendungen auf die Lehre notwendig ist.

Die bisherigen Literaturrecherchen zum Einsatz von KI in der Hochschulbildung sind überwiegend auf technische Anwendungen wie automatische Bewertungssysteme, intelligente Tutorensysteme, Vorhersagemodelle für akademische Leistungen sowie die Personalisierung von Lernressourcen konzentriert. Ethische Aspekte wie Datenschutz und Bias werden zwar thematisiert, jedoch oft isoliert von den technischen Analysen betrachtet. Gleichzeitig werden eine stärkere pädagogische Einbindung sowie weiterführende empirische Forschung zur langfristigen Wirkung von KI-Anwendungen gefordert.

Folgende Forschungslücken bei systematischen Literaturrecherchen konnten identifiziert werden:

- **Integrative Perspektive:** Es fehlen Studien, die technische und ethische Dimensionen von KI im Bildungswesen ganzheitlich vereinen.
- **Bewältigung spezifischer Herausforderungen:** Detaillierte Analysen zur Überwindung von Problemen wie ‚Halluzinationen‘ großer Sprachmodelle (LLMs) und der mangelnden Erklärbarkeit komplexer KI-Modelle sind bislang unzureichend.
- **Praktische Datenschutzlösungen:** Konkrete Handlungsempfehlungen für Bildungseinrichtungen zur effektiven Implementierung von Datenschutzmaßnahmen fehlen.
- **Vermeidung von Bias:** Es mangelt an spezifischen Strategien und Empfehlungen zur Reduzierung von Bias in KI-Systemen.

Diese Forschungslücken verdeutlichen den dringenden Bedarf an umfassenderen und integrativen Studien, die sowohl technische als auch ethische Aspekte berücksichtigen. Insbesondere ist es notwendig, praxisnahe Lösungen zu entwickeln, die Bildungseinrichtungen dabei unterstützen, KI-Systeme verantwortungsvoll und effektiv zu implementieren. Ohne eine ganzheitliche Betrachtung besteht die Gefahr, dass technische Fortschritte unkontrolliert voranschreiten und dabei ethische Standards vernachlässigt werden, was langfristig das Vertrauen in KI-Anwendungen im Bildungsbereich untergraben könnte.

Daraus lässt sich folgende Forschungsfrage ableiten.

Forschungsfrage: Wie können KI-Systeme, insbesondere Large Language Models (LLMs), im akademischen und hochschulischen Kontext konzipiert und implementiert werden, um sowohl

die Lernprozesse zu verbessern und technische Effizienz zu gewährleisten als auch ethische Standards wie Datenschutz einzuhalten?

Ziele dieser Arbeit:

1. **Systematische Analyse:** Untersuchung der aktuellen Potenziale und Herausforderungen von KI und LLMs im Bildungsbereich, um ein ganzheitliches Verständnis des Themas zu entwickeln.
2. **Identifikation von Anforderungen:** Bestimmung der spezifischen technischen und ethischen Anforderungen, die bei der Gestaltung von KI-Systemen für den Bildungsbereich berücksichtigt werden müssen.
3. **Entwicklung von Empfehlungen für den Einsatz von KI-Systemen an Hochschulen:** Erarbeitung konkreter Empfehlungen für Bildungseinrichtungen, die eine verantwortungsvolle und effektive Implementierung von KI-Technologien ermöglichen.

Durch die Verknüpfung technischer und ethischer Perspektiven wird mit dieser Arbeit darauf abgezielt, eine praktische Orientierung für die Gestaltung und die Nutzung von KI-Systemen im Bildungsbereich zu bieten.

Methodik

Um den aktuellen Forschungsstand zur Integration von KI und LLMs im Bildungsbereich zu erfassen, wurde eine systematische Literaturrecherche gemäß den PRISMA-Richtlinien (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA, 2020) durchgeführt. Die Recherche erfolgte im September 2024 in den wissenschaftlichen Datenbanken Web of Science, Scopus und Google Scholar. Der Zweck dieser systematischen Übersichtsarbeit besteht darin, die vorgestellte Forschungsfrage basierend auf einer klaren, systematischen und reproduzierbaren Suchstrategie zu beantworten (Natukunda, & Muchene, 2023). Dabei werden Ein- und Ausschlusskriterien verwendet, um relevante Studien zu identifizieren (Gough et al., 2017).

Die Suche wurde sowohl in deutscher als auch in englischer Sprache durchgeführt, um ein breites Spektrum an relevanter Literatur abzudecken. Für jede Datenbank wurden spezifische Suchstrings entwickelt, die Boolesche Operatoren und Wildcards enthielten. Die genauen Suchstrings und die Anzahl der gefundenen Studien pro Datenbank sind in Tabelle 1 dargestellt.

Tabelle 1: Suchstrings und Anzahl der gefundenen Studien pro Datenbank

Datenbank	Suchstring	Anzahl gefundener Studien
Web of Science	<p>Deutsch: TS=(„Künstliche Intelligenz“ AND Bildung) OR („Künstliche Intelligenz“ AND Hochschulbildung) OR („Large Language Models“ AND Bildung) OR („LLMs“ AND „ethische Implikationen“ AND Bildung) OR („KI“ AND Datenschutz AND Bildung)</p>	845
	<p>Englisch: TS=(„Artificial Intelligence“ AND Education) OR („Artificial Intelligence“ AND „Higher Education“) OR („Large Language Models“ AND Education) OR („LLMs“ AND „Ethical Implications“ AND Education) OR („AI“ AND „Data Privacy“ AND Education) OR („LLMs“ AND „Ethical Implications“ AND Education) OR („AI“ AND „Energy Ressources“ AND Education)</p>	
Scopus	<p>Deutsch: TITLE-ABS-KEY(„Künstliche Intelligenz“ AND Bildung) OR („Künstliche Intelligenz“ AND Hochschulbildung) OR („Large Language Models“ AND Bildung) OR („LLMs“ AND „ethische Implikationen“ AND Bildung) OR („KI“ AND Datenschutz AND Bildung)</p>	367
	<p>Englisch: TITLE-ABS-KEY(„Artificial Intelligence“ AND Education) OR („Artificial Intelligence“ AND „Higher Education“) OR („Large Language Models“ AND Education) OR („LLMs“ AND „Ethical Implications“ AND Education) OR („AI“ AND „Data Privacy“ AND Education)</p>	
Google Scholar	<p>Deutsch & Englisch: Alle oben genannten Suchstrings wurden eingegeben, ergänzt durch Wildcards () zur Abdeckung verschiedener Wortendungen (z. B. „educat“ für „education“, „educational“ etc.).</p>	734

Insgesamt wurden 1946 Studien identifiziert (Web of Science: 845, Scopus: 367, Google Scholar: 734). Nach dem Entfernen von Duplikaten verblieben 467 einzigartige Studien für die weitere Analyse.

Einschluss- und Ausschlusskriterien

Um die Relevanz und die Qualität der ausgewählten Studien sicherzustellen, wurden klare Einschluss- und Ausschlusskriterien festgelegt:

- **Einschlusskriterien:**
 - **Zeitlicher Rahmen:** Veröffentlichungen zwischen 2014 und 2024, um die Aktualität der Forschung zu gewährleisten.
 - **Publikationstyp:** Publikationen mit Blind Review, Konferenzbeiträge und Bücher.
 - **Sprache:** Veröffentlichungen in deutscher und englischer Sprache.
 - **Inhaltliche Relevanz:** Studien, die sich explizit mit dem Einsatz von KI und LLMs im Bildungsbereich oder mit Implikationen auf diesen Bereich befassen und empirische Daten oder theoretische Beiträge liefern.

- **Ausschlusskriterien:**
 - **Themenfremde Studien:** Arbeiten, die sich nicht direkt auf den Bildungsbereich beziehen oder KI in einem anderen Kontext behandeln.
 - **Mangelnde Verfügbarkeit:** Veröffentlichungen ohne zugänglichen Volltext sowie unvollständige Arbeiten, außer sie erschienen relevant für die Arbeit (12 Artikel wurden angekauft).
 - **Qualitative Mängel:** Studien mit unzureichender methodischer Qualität (z. B. nicht beschriebene Stichprobe, fehlende oder ungenaue Beschreibung der Methoden) oder fehlender wissenschaftlicher Evidenz.

Für die Aufarbeitung der Literatur wurde das KI-Management-System der Privaten Pädagogischen Hochschule Burgenland genutzt. Folgende Arbeitsschritte wurden mit Unterstützung des KI-Systems durchgeführt: (1) Identifikation von Duplikaten (KI@PPHBurgenland, 2024a), (2) Zusammenfassung von Abstracts und Titeln zur Identifikation relevanter Studien (KI@PPHBurgenland, 2024b) und (3) Untersuchung und Screening der Abstracts nach KI-Anwendungen im Bildungsbereich (KI@PPHBurgenland; LLM OpenAI, 2024c) (Prompting und Outputvariationen im Literaturverzeichnis).

Diese von der KI@PPHB durchgeführten Arbeitsschritte wurden allesamt händisch nachkontrolliert und anschließend dem KI-System zu einer erneuten Kontrolle übergeben.

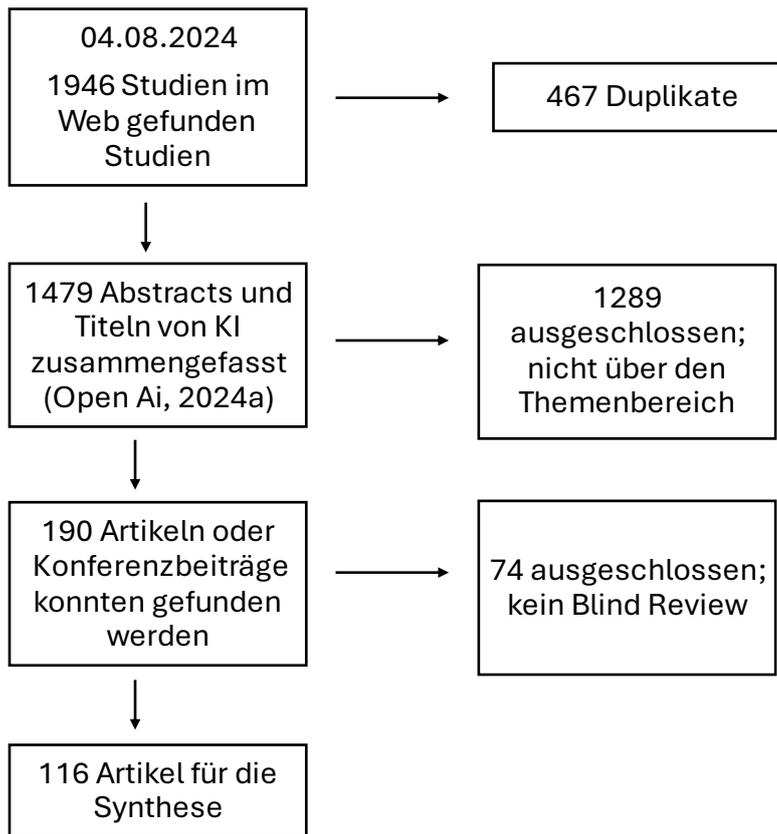


Abbildung 1: PRISMA diagram (slightly modified)

Analyseverfahren

Das Kodierungsschema wurde basierend auf den Forschungsfragen und unter Einbeziehung bestehender Literatur entwickelt (Mayring, 2015; 2019). Der Prozess erfolgte in drei Phasen:

- **Offenes Kodieren:** In dieser Phase wurden relevante Textpassagen aus den Volltexten der ausgewählten Studien identifiziert und mit initialen Codes versehen. Beispielsweise wurde ein Abschnitt, der die Anpassung von Lernmaterialien an individuelle Bedürfnisse beschreibt, mit dem Code „Personalisierung“ gekennzeichnet.
- **Axiales Kodieren:** Die initialen Codes wurden anschließend zu übergeordneten Kategorien zusammengefasst. Diese Phase diente der Strukturierung und der Verknüpfung der Codes, um größere Zusammenhänge und Muster zu erkennen.
- **Selektives Kodieren:** Abschließend wurden die Kategorien zu einem Gesamtmodell integriert, das die Forschungsfrage umfassend adressiert. Dies beinhaltete die Beziehung zwischen den Kategorien „Potenziale von KI“, „Herausforderungen bei der Implementierung“, „Ethische Implikationen“ sowie „Vorgeschlagene Maßnahmen und Strategien“.

Definition und Operationalisierung der Hauptkategorien und Codes

Für jede Hauptkategorie wurden spezifische Unterkategorien und Codes definiert, um eine präzise Erfassung der relevanten Daten zu ermöglichen. Im Folgenden werden die zentralen Kategorien und deren Operationalisierung dargestellt:

1. Potenziale von KI und LLMs im Bildungsbereich

Diese Hauptkategorie umfasst die vielfältigen Möglichkeiten, die der Einsatz von KI und LLMs im Bildungssektor bietet. Die Subkategorien unterteilen diese Potenziale in spezifische Bereiche.

- **1.1 Personalisierung des Lernens**
 - Systeme, die Lerninhalte basierend auf den individuellen Fortschritten und Bedürfnissen der Lernenden anpassen.
 - Erstellung maßgeschneiderter Lernwege, die auf den Fähigkeiten und den Zielen der Studierenden basieren.
- **1.2 Effizienzsteigerung**
 - Einsatz von KI zur Automatisierung von Verwaltungsprozessen wie Kursplanung und Studierendenverwaltung.
 - Nutzung von LLMs zur automatisierten Bewertung von Prüfungen und zur Generierung von personalisiertem Feedback.
- **1.3 Innovative Lernmethoden**
 - Integration von Chatbots und LLMs zur Bereitstellung personalisierter Unterstützung und unmittelbarer Feedbacks.

2. Herausforderungen bei der Implementierung von KI

Diese Hauptkategorie identifiziert die wesentlichen Hindernisse und Schwierigkeiten, die bei der Einführung von KI-Technologien im Bildungsbereich auftreten können.

- **2.1 Datenqualität**
 - Sicherstellung der Genauigkeit, der Vollständigkeit und der Konsistenz der verwendeten Daten sowie Zugang zu hochwertigen Datensätzen.
- **2.3 Ressourcen und Infrastruktur**
 - Hohe Investitionen für die Entwicklung und die Implementierung von KI-Systemen.
 - Hoher Energieverbrauch und ökologische Auswirkungen durch die Nutzung großer KI-Modelle.

3. Ethische Implikationen

Diese Hauptkategorie untersucht die ethischen Implikationen, die mit dem Einsatz von KI und LLMs im Bildungsbereich einhergehen.

- **3.1 Datenschutz**
 - Sicherstellung, dass KI-Systeme den gesetzlichen Datenschutzanforderungen entsprechen.
 - Maßnahmen zum Schutz der Daten der Studierenden vor unbefugtem Zugriff.
- **3.2 Bias – systematische Verzerrungen in den Entscheidungen oder Ergebnissen von KI-Systemen**
 - Maßnahmen zur Identifikation und Reduktion von Bias in KI-Modellen.

4. Empfohlene Maßnahmen und Strategien

Diese Hauptkategorie umfasst konkrete Handlungsempfehlungen und Strategien, die Bildungseinrichtungen ergreifen sollten, um die Implementierung von KI-Technologien verantwortungsvoll und effektiv zu gestalten.

- **4.1 Leitlinien**
 - Erstellung von Richtlinien, die den verantwortungsvollen Einsatz von KI im Bildungsbereich regeln.
- **4.2 Professionalisierung**
 - Programme zur Weiterbildung von Lehrkräften im Umgang mit KI-Technologien.
 - Schulungsprogramme zur Förderung des Bewusstseins für Datenschutz, Fairness und Verantwortlichkeit.

Beispiele zur Operationalisierung

Um die Operationalisierung der Codes zu veranschaulichen, werden hier zwei konkrete Beispiele aus den analysierten Studien dargestellt:

- **Beispiel 1: Personalisierung des Lernens**
 - **Textpassage:** „Das adaptive Lernsystem analysiert kontinuierlich die Fortschritte der Studierenden und passt die Schwierigkeitsgrade der Aufgaben entsprechend an.“
 - **Angewandter Code:** „Adaptive Lernsysteme“
 - **Operationalisierung:** Diese Passage beschreibt die dynamische Anpassung der Lerninhalte basierend auf den individuellen Fortschritten der Studierenden, was direkt dem Code „Adaptive Lernsysteme“ zugeordnet wird.

• **Beispiel 2: Datenschutz und Privatsphäre**

- **Textpassage:** „Die Implementierung von Verschlüsselungstechnologien gewährleistet, dass personenbezogene Daten der Studierenden vor unbefugtem Zugriff geschützt sind.“
- **Angewandter Code:** „Datenschutz und Privatsphäre“
- **Operationalisierung:** Die Passage beschreibt konkrete Maßnahmen zur Sicherung personenbezogener Daten, was dem Code „Datenschutz und Privatsphäre“ entspricht.

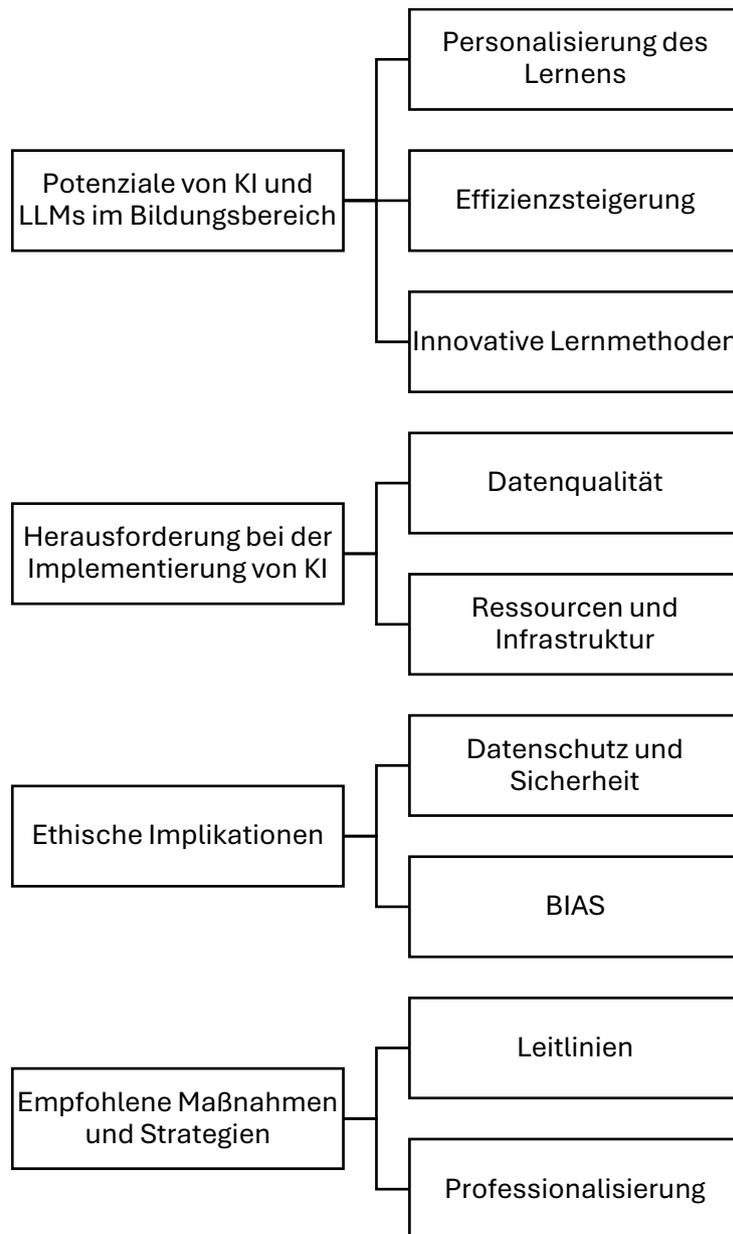


Abbildung 2: Schematische Darstellung des Kodierungsschemas

Datenextraktion und -synthese

- **Datenextraktion:** Die codierten Daten wurden extrahiert und in einer Tabelle festgehalten, die für jede Studie die relevanten Kategorien und Schlüsselbefunde enthält.
- **Qualitative Inhaltsanalyse:** Eine qualitative Inhaltsanalyse nach Mayring (2015) wurde durchgeführt, um Muster, Zusammenhänge und zentrale Aussagen zu identifizieren.
- **Integration in den Forschungsstand:** Zitate und Schlüsselbefunde wurden gesammelt und in die Synthese des Forschungsstands integriert, um die Argumentation zu untermauern.

Aufarbeitung der identifizierten Studien

Im vorliegenden Kapitel soll anhand der Operationalisierung der Hauptkategorien eine Aufbereitung des Forschungsstands erfolgen.

1. Potenziale von KI und LLMs im Bildungsbereich

1.1 Personalisierung des Lernens

Künstliche Intelligenz und LLMs revolutionieren das Bildungswesen, indem sie adaptive Lernumgebungen schaffen, die personalisierte Lernpfade für Studierende ermöglichen. Ng und Fung (2024) zeigen, dass LLMs wie GPT-4 durch die Integration von Prompt-Engineering präzise und maßgeschneiderte Lernpfade für Lernende erstellen können, was zu einer höheren Lerneffizienz und verbesserten Ergebnissen führt. Adaptive Lernsysteme profitieren ebenfalls von LLMs, da sie die Interaktion zwischen Mensch und Maschine verbessern und personalisierte Rückmeldungen in Echtzeit ermöglichen, was zu einer flexibleren und dynamischeren Lernerfahrung führt (Wen et al., 2024). Diese Technologien tragen zur individuellen Lernpfadplanung bei, indem sie Lernstrategien dynamisch anpassen und dadurch den Lernprozess für den Einzelnen optimieren.

Während LLMs einerseits das Potenzial haben, eine tiefere Personalisierung des Lernens zu fördern, müssen dabei andererseits Herausforderungen wie die Vermeidung von Fehlern und ethische Bedenken hinsichtlich Verzerrungen berücksichtigt werden. Suresh und Misra (2024) betonen, dass die Genauigkeit der von LLMs generierten Inhalte sowie die Vermeidung von algorithmischen Verzerrungen entscheidend für den erfolgreichen Einsatz im Bildungsbereich sind.

Zawacki-Richter et al. (2019) ergänzen in ihrer systematischen Literaturübersicht von 146 Studien, dass solche Technologien individuelle Lernpfade ermöglichen, die auf die spezifischen Bedürfnisse der Lernenden zugeschnitten sind. Modran et al. (2024) identifizieren die Begrenzungen traditioneller Tutoring-Methoden sowie die Schwächen des alleinigen Einsatzes von LLMs wie ChatGPT in der Hochschulbildung. Sie entwickeln ein innovatives intelligentes Chatbot-Tutoring-System, das den Retrieval-Augmented-Generation(RAG)-Ansatz mit einem maßgeschneiderten LLM kombiniert, um präzise, kontextuell relevante und personalisierte Unterstützung zu bieten. Ziel des Systems ist es, das Verständnis und das Engagement der Studierenden

zu verbessern sowie eine bereicherte Bildungserfahrung und gesteigerte akademische Leistungen zu ermöglichen. Solche Systeme haben das Potenzial, herkömmliche Lernansätze zu verbessern und eine tiefere Einbindung der Lernenden zu erreichen, da sie in Echtzeit auf den Fortschritt und die Anforderungen der Nutzer reagieren können. Akinwalere und Ivanov (2022) betonen, dass KI im Bildungsbereich erhebliche Chancen bietet, insbesondere durch Anwendungen wie Profilierung, Vorhersage, intelligente Tutorensysteme und personalisierte Lernansätze.

1.2 Effizienzsteigerung

Die Integration von KI in Bildungssysteme eröffnet zahlreiche Möglichkeiten zur Verbesserung des Lernens, indem personalisierte Lernpfade geschaffen und administrative Aufgaben automatisiert werden. Abimbola et al. (2024) heben hervor, dass KI das Potenzial hat, Lernprozesse an individuelle Bedürfnisse anzupassen und somit die Lernerfolge zu steigern. Gleichzeitig sind jedoch ethische Herausforderungen wie der Datenschutz und die Gefahr algorithmischer Verzerrungen zu beachten, da der Umgang mit großen Datenmengen sensible Schülerdaten betrifft und ungleiche Zugänge zu diesen Technologien bestehende Bildungsungleichheiten verschärfen können.

Erweiterte Studien zeigen, dass LLMs die automatisierte Bewertung weiter revolutionieren. Larrondo et al. (2024) setzten LLMs für das automatisierte Feedback bei komplexen technischen Aufgaben ein und erzielten eine hohe Genauigkeit von 80 % bei der Klassifikation von Problemstellungen. Stamper et al. (2024) unterstreichen, dass LLMs in intelligenten Tutorensystemen eingesetzt werden können, um personalisiertes Feedback zu generieren, was die Effizienz der Lehrkräfte erheblich steigert und die Lernleistung der Studierenden verbessert.

Weitere Forschungen wie die von Estévez-Ayres et al. (2024) haben gezeigt, dass LLM-basierte Tools typische Fehler in Programmieraufgaben, z. B. Deadlocks, identifizieren und konsistentes Feedback geben können. Obwohl die Genauigkeit bei der Fehlererkennung noch ausbaufähig ist, zeigen diese Systeme vielversprechende Ergebnisse in der Automatisierung von Bewertungsprozessen.

1.3 Innovative Lehrmethoden

Der Einsatz von Chatbots und LLMs im Bildungsbereich bietet vielversprechende Möglichkeiten. Deng und Yu (2023) untersuchten in einer systematischen Literaturrecherche 32 wissenschaftliche Studien zwischen 2010 und 2022 mit insgesamt 2201 Proband:innen den Einsatz von Chatbot-Technologie im Bildungsbereich. Sie stellten fest, dass Chatbots das Lerninteresse und die Lernleistung fördern können. Allerdings wiesen sie auch auf ethische und technologische Herausforderungen hin, insbesondere in Bezug auf Datenschutz und die Vermeidung algorithmischer Verzerrungen.

Winkler und Söllner (2018) analysierten in ihrer State-of-the-Art-Analyse den Einsatz von Chatbots in der Bildung und betonten deren Potenzial, personalisierte Unterstützung und unmittelbares Feedback zu bieten. In einer Studie von Fryer, Nakao und Thompson (2019) wird der Einsatz von Chatbots als Gesprächspartner im Sprachlernen untersucht. Dabei führten Studierende Gespräche sowohl mit menschlichen Partner:innen als auch mit Chatbots und bewerteten anschließend ihr Interesse sowie ihren Lernerfolg. Die Ergebnisse zeigen, dass das vorherige Interesse an menschlichen Gesprächspartner:innen ein starker Prädiktor für das Interesse an

Chatbot-Gesprächen ist. Außerdem ist die Sprachkompetenz stärker mit dem Interesse an Chatbots als mit dem an menschlichen Partner:innen verknüpft. Zudem erhöht das Gefühl, mit Chatbots mehr gelernt zu haben, das Interesse an den Lernaufgaben, selbst bei auftretenden Kommunikationsschwierigkeiten. Die Studie liefert wertvolle Hinweise für den effektiven Einsatz und die Weiterentwicklung von Chatbots im Sprachlernkontext. Ein praktisches Beispiel ist der KI-gestützte Chatbot ‚Jill Watson‘ an der Georgia Tech University, der Studierende bei häufig gestellten Fragen unterstützt und Lehrkräfte entlastet (Taneja et al., 2024).

Akpan et al. (2024) untersuchten in ihrer umfassenden Analyse die Rolle von generativer KI und Chatbots im Bildungswesen. Die Studie zeigt, dass Chatbots in verschiedenen Disziplinen wie Informatik, Gesundheitswesen und Ingenieurwesen eingesetzt werden, um die Lernumgebung zu verbessern und automatisiertes Feedback zu ermöglichen. Allerdings weisen die Autor:innen auch auf potenzielle Nachteile wie den Verlust menschlicher Interaktionen und die Notwendigkeit kritischen Denkens hin (Duran, 2024). Zudem zeigen Ali et al. (2024), dass Chatbots wie Google Bard und Microsoft Bing Chat dabei helfen, akademische Kompetenzen zu stärken, insbesondere bei der Quellenbewertung und im Bereich kritischen Denkens.

2. Herausforderungen bei der Implementierung von KI

2.1 Datenqualität

Die Nutzung von Big Data im Bildungswesen muss strenge Rahmenbedingungen für Datenqualität und Datenschutz einhalten. Studien wie die von Budach et al. (2022) heben hervor, dass Datenqualität in mehreren Dimensionen – wie Genauigkeit, Vollständigkeit und Konsistenz – eine Schlüsselrolle für die Leistung von Machine-Learning-Modellen spielt. Fehlerhafte oder unvollständige Daten verringern das Vertrauen in KI-Systeme durch unzuverlässige Vorhersagen. Darüber hinaus bestehen in vielen Bereichen Herausforderungen hinsichtlich der Verfügbarkeit von Daten, insbesondere bei eingeschränktem Zugang zu hochwertigen Datensätzen.

2.3 Ressourcen und Infrastruktur

Die Entwicklung und die Implementierung von KI-Systemen erfordern erhebliche finanzielle Investitionen. Die Studie von Li (2023) zeigt, dass der Einsatz von KI im Bildungswesen bestehende Ungleichheiten verstärken kann, da ressourcenarme Institutionen Schwierigkeiten haben, die notwendigen finanziellen Mittel für die Implementierung von KI-Systemen bereitzustellen. Diese Diskrepanz führt dazu, dass wohlhabendere Institutionen technologisch weiter fortschreiten, während andere zurückbleiben. Li betont, dass internationale Zusammenarbeit und eine gerechtere Verteilungen von KI-Ressourcen notwendig sind, um diese Ungleichheiten zu überwinden und die Chancengleichheit im Bildungsbereich zu fördern.

In einer Studie von Gowda et al. (2023) wurde der steigende Energieverbrauch von Deep-Learning-Modellen und deren Auswirkungen auf die Umwelt untersucht. Im Gegensatz zu bestehenden Forschungsarbeiten, die primär die Modellgenauigkeit fokussieren, analysierten die Autor:innen das Verhältnis von Genauigkeit und Energieverbrauch. Diese Forschung erweitert das Verständnis für die energieeffiziente Entwicklung von KI-Modellen und betont die Notwendigkeit, ökologische Nachhaltigkeit in der Modellarchitektur zu berücksichtigen.

Strubell et al. (2019) untersuchten den erheblichen Energie- und Ressourcenverbrauch, der mit dem Training von LLMs einhergeht. Sie zeigten, dass Modelle wie BERT und GPT-2 nicht nur hohe finanzielle Kosten verursachen, sondern auch einen erheblichen ökologischen Fußabdruck hinterlassen. Ihre Forschung verdeutlicht, dass der Zugang zu leistungsfähiger Hardware oft auf große Forschungseinrichtungen und Unternehmen beschränkt ist, was zu einer Ungleichheit in der Forschungsgemeinschaft führt. Daher betonen die Autor:innen die Notwendigkeit, effizientere Algorithmen und Hardwarelösungen zu entwickeln sowie den Zugang zu Rechenressourcen gerechter zu gestalten.

In ihrer Studie untersuchten Hidalgo et al. (2023) ebenfalls die steigenden Umweltbelastungen im Zuge des Energieverbrauchs von KI-Systemen und zeigten auf, dass diese bis zu 9 % des weltweiten Strombedarfs ausmachen. Die Autoren betonen die Notwendigkeit, KI-Modelle energieeffizienter zu gestalten, um CO₂-Emissionen zu reduzieren. Sie schlagen einen ganzheitlichen Ansatz vor, der Optimierungen auf allen Ebenen der KI-Entwicklung einbezieht. Ziel ist es, die Leistungsfähigkeit von KI zu erhalten, während der Energieverbrauch gesenkt wird, um so die Nachhaltigkeit von KI-Systemen zu fördern.

3. Ethische Implikationen

3.1. Datenschutz und Sicherheit

Die Einhaltung gesetzlicher Vorgaben wie der Datenschutz-Grundverordnung (DSGVO) stellt Bildungseinrichtungen vor erhebliche Herausforderungen bei der Implementierung von KI-Systemen. Der Schutz personenbezogener Daten ist von entscheidender Bedeutung, um das Vertrauen der Studierenden in KI-Anwendungen aufrechtzuerhalten. Drachler und Greller (2016) argumentieren, dass unzureichende Datenschutzmaßnahmen das Vertrauen in solche Technologien untergraben können, was die Akzeptanz von KI im Bildungsbereich negativ beeinflussen könnte. Bai et al. (2024) ergänzen, dass die Einhaltung der DSGVO sowie die Implementierung von Datenschutzmechanismen angesichts neuer KI-Technologien wie ChatGPT zunehmend herausfordernd werden. Solche Systeme verarbeiten umfangreiche personenbezogene Daten.

Die Verarbeitung großer Mengen sensibler Daten wirft erhebliche Datenschutzbedenken auf. Slade und Prinsloo (2013) diskutierten bereits früh die ethischen Implikationen von Lernanalysen und betonten die Notwendigkeit eines ausgewogenen Ansatzes, der den Nutzen solcher Technologien mit den potenziellen Risiken in Einklang bringt.

Besonders in Bereichen wie Bildung und Gesundheitswesen, in denen sensible personenbezogene Daten verarbeitet werden, stellen Datenschutzverletzungen eine kritische Bedrohung dar. KI-Systeme, einschließlich maschinellen Lernens und automatisierter Datenerfassungssysteme, basieren häufig auf großen Datenmengen, was das Risiko von Datenlecks und unsachgemäßer Verarbeitung erhöht. In der Gesundheitsbranche werden daher Maßnahmen wie Verschlüsselung und Anonymisierung empfohlen (Singhal, 2024).

Algorithmische Verzerrungen können unbewusst durch KI-Anwendungen verstärkt werden, was insbesondere in sensiblen Bereichen wie der Bildung zu Ungerechtigkeiten führen kann. Die Studie von Omughelli et al. (2024) thematisiert die Rolle von KI im Bildungswesen mit Fokus auf

Bias in der Vorhersage von Studienerfolgen. Mittels Analyse des *Open University Learning Analytics Datasets* und des Einsatzes von KI-Algorithmen wurden Verzerrungen identifiziert, die zu ungleichen Ergebnissen für verschiedene demografische Gruppen führen könnten. Die Autor:innen betonen daher, dass es notwendig ist, ethische Richtlinien und Techniken zur Bias-Erkennung zu entwickeln, um eine faire und inklusive Lernumgebung zu gewährleisten.

3.2. BIAS – systematische Verzerrungen in den Entscheidungen oder Ergebnissen von KI-Systemen

Algorithmische Verzerrungen in KI-Systemen stellen eine erhebliche Herausforderung dar, da sie diskriminierende Ergebnisse erzeugen und soziale Ungleichheiten verstärken können. Baker und Hawn (2022) unterstreichen die Notwendigkeit, Bias in KI-Modellen zu identifizieren und zu minimieren, um gerechte Ergebnisse im Bildungswesen zu gewährleisten. Mit Studien wie der von Pagano et al. (2023) wird gezeigt, dass die Verwendung von Metriken wie ‚Equalized Odds‘ und ‚Demographic Parity‘ helfen kann, Verzerrungen zu erkennen und zu mindern. Diese Ansätze sind entscheidend, um Chancengleichheit für alle Lernenden zu fördern und sicherzustellen, dass KI-Systeme keine unbewussten Vorurteile reproduzieren oder verstärken.

Floridi und Cowls (2019) untersuchten in ihrem Artikel die Vielzahl existierender ethischer Prinzipien für KI und identifizierten eine signifikante Übereinstimmung zwischen diesen Ansätzen. Basierend auf ihrer Analyse entwickelten sie ein einheitliches Rahmenwerk, das aus fünf Kernprinzipien besteht: Transparenz, Gerechtigkeit, Verantwortung, Datenschutz und menschliche Kontrolle. Dieses Rahmenwerk soll als Grundlage für die Schaffung konsistenter Gesetze, technischer Standards und Best Practices dienen, um eine ethisch verantwortungsvolle und gesellschaftlich nützliche KI-Entwicklung zu fördern.

Holmes et al. (2019) argumentieren, dass KI-Systeme so gestaltet sein sollten, dass sie die Chancengleichheit fördern, indem sie allen Lernenden unabhängig von deren Hintergrund Zugang zu hochwertigen Bildungsressourcen bieten. Dies kann durch die Entwicklung geeigneter Algorithmen und die Vermeidung diskriminierender Entscheidungen erreicht werden, wie Misino et al. (2024) betonen. Die Förderung von Chancengleichheit durch KI im Bildungswesen erfordert somit eine bewusste Gestaltung und die kontinuierliche Überprüfung der eingesetzten Technologien, um sicherzustellen, dass diese inklusiv und gerecht sind.

Das Problem algorithmischer Verzerrungen ist eng mit der ‚Black-Box‘-Natur³ vieler KI-Algorithmen verknüpft, was die Nachvollziehbarkeit von Entscheidungsprozessen behindert und die Erkennung von Verzerrungen erschwert. Studien wie die von Chinta et al. (2024) zeigen, dass Vorurteile in den Daten oder durch algorithmische Schwächen zu Ungleichheiten führen können, die insbesondere benachteiligte Gruppen negativ beeinflussen. Aspekte wie Datenerhebung, Vorverarbeitung und algorithmisches Design spielen eine entscheidende Rolle für die Sicherstellung von Gerechtigkeit. Chaudhary (2024) hebt außerdem die Bedeutung von Transparenz in KI-Systemen mittels erklärbarer KI-Modellen (XAI) hervor, um Vertrauen zu schaffen und ethische Bedenken zu adressieren. Der Mangel an Algorithmustransparenz erschwert es, Verzerrungen zu erkennen, was in Bereichen wie der Bildung zu Ungerechtigkeiten bei der Bewertung und der Auswahl führen kann (Samek et al., 2017). Die Forderung nach algorithmischer

³ Eine „Black Box“ bezeichnet in der Informatik und speziell im Kontext von Künstlicher Intelligenz (KI) ein System, dessen interne Funktionsweise und Entscheidungsprozesse für den Anwender oder Außenstehenden nicht transparent sind (von Eschenbach & Warren, 2021).

Fairness ist daher zentral, um sicherzustellen, dass KI-Systeme Entscheidungen transparent und gerecht treffen.

4. Empfohlene Maßnahmen und Strategien

4.1. Entwicklung ethischer Leitlinien

Die rasche Verbreitung von generativer künstlicher Intelligenz (GenAI) in Hochschulen seit der Einführung von ChatGPT im November 2022 hat die Notwendigkeit für klare ethische Leitlinien und Richtlinien hervorgehoben. Verschiedene Studien und Initiativen beschäftigen sich mit den Herausforderungen und Chancen, die der Einsatz von GenAI in der Hochschulbildung mit sich bringt.

McDonald et al. (2024) untersuchten beispielsweise 116 US-Universitäten und fanden heraus, dass 63 % dieser Einrichtungen die Nutzung von GenAI fördern, wobei viele detaillierte Anweisungen für den Einsatz im Unterricht bieten, insbesondere im Bereich des Schreibens. In diesem Kontext werden häufig Themen wie Datenschutz, ethische Fragen und Vielfalt diskutiert. Die Autor:innen warnen außerdem davor, dass die umfangreichen Anpassungen der Lehrmethoden die Lehrkräfte stark belasten könnten.

In einer ähnlichen Richtung identifizieren Ioku et al. (2024) vier Profile von Universitäten basierend auf ihrer Akzeptanz generativer KI wie ChatGPT. Die Ergebnisse betonen, dass klare Richtlinien und unterstützende Umgebungen entscheidend für den verantwortungsvollen Einsatz von KI in der Hochschulbildung sind. Ergänzend dazu untersucht Sanasintani (2023) die Integration von KI in Hochschulcurricula, um auf globale Veränderungen zu reagieren. In einer quantitativen Umfrage unter 20 Studierenden gaben 70 % an, mit KI vertraut zu sein, und 80 % glauben, dass KI die Qualität des Lernens verbessern kann, während sich 20 % neutral äußerten.

Um diesen Entwicklungen gerecht zu werden, hat der PH-Verbund Süd-Ost Leitlinien für die Nutzung von KI im Hochschulbereich veröffentlicht (Leitgeb et al., 2024). Diese geben einen umfassenden Rahmen für den Einsatz von KI in der Lehre und bei Prüfungsbewertungen vor. Die Richtlinien umfassen rechtliche Aspekte, IT-Sicherheit und fachdidaktische Empfehlungen. Generative KI-Anwendungen dürfen demnach in der Lehre unter bestimmten Voraussetzungen verwendet werden, wobei rechtliche Vorgaben und Datenschutz eingehalten werden müssen. Die Handlungsempfehlungen beinhalten, wie KI verantwortungsvoll und ethisch in Lehrkontexte integriert werden kann, wobei die Eigenleistung der Studierenden sichergestellt werden soll. Es wird empfohlen, klare Regelungen für den Einsatz und die Bewertung KI-generierter Arbeiten zu schaffen (Leitgeb et al., 2024).

4.2. Professionalisierung

Die Schulung von Lehrkräften im Umgang mit KI-Technologien ist entscheidend, um deren Potenziale effektiv in den Unterricht zu integrieren. Die Studie von Umanets et al. (2024) identifizierte die Entwicklung technischer, pädagogischer sowie ethischer und rechtlicher Kompetenzen bei Lehrkräften als notwendig, um KI-Technologien erfolgreich im Bildungsprozess zu nutzen. Dies umfasst die Fähigkeit, KI-Tools effektiv einzusetzen, um den Unterricht zu verbessern

und personalisierte Lernwege für Studierende zu ermöglichen. Weiterhin zeigen Brandão et al. (2024), dass die Integration dieser Kompetenzen in Weiterbildungsprogramme essenziell ist, um Lehrkräfte auf die Herausforderungen und die Möglichkeiten der KI-Nutzung vorzubereiten. Weiterbildungsprogramme sollten zudem ethische Aspekte adressieren, um das Bewusstsein für Datenschutz, Fairness und Verantwortlichkeit zu schärfen. Zawacki-Richter et al. (2019) betonen die Notwendigkeit, Pädagog:innen in ethischen Fragen der KI-Nutzung zu schulen. Brandão et al. (2024) unterstreichen ebenfalls, dass ethische Aspekte wie Datenschutz und Fairness integrale Bestandteile von Weiterbildungsprogrammen sein sollten, um Lehrkräfte in der sicheren und ethischen Nutzung von KI zu schulen. Eine ethische Sensibilisierung ist entscheidend, um die Risiken und die Herausforderungen der KI-Nutzung zu verstehen und verantwortungsvoll mit den entsprechenden Daten umzugehen. Dies fördert nicht nur den verantwortungsbewussten Einsatz von KI-Technologien, sondern stärkt auch das Vertrauen der Lernenden in die verwendeten Systeme. Durch die Einbindung ethischer Schulungen können Bildungseinrichtungen sicherstellen, dass Lehrkräfte nicht nur technisch kompetent, sondern auch ethisch sensibilisiert sind, was zur Schaffung eines fairen und transparenten Bildungsumfelds beiträgt.

Synthese der aktuellen Forschung zu Potenzialen, Herausforderungen und Strategien für KI-Systeme in Hochschulen

Durch Anwendung des axialen Kodierens wurden die in der Literatur identifizierten Potenziale, Herausforderungen, ethischen Implikationen und empfohlenen Maßnahmen im Kontext der Implementierung von KI und LLMs im Hochschulbereich systematisch analysiert und in Beziehung gesetzt. Die aktuelle Literatur betont das erhebliche Potenzial von KI und insbesondere von LLMs im Hochschulbereich, weist jedoch zugleich auf bedeutende Herausforderungen und ethische Implikationen hin. Künstliche Intelligenz ermöglicht die Personalisierung des Lernens durch adaptive Lernumgebungen und individuelle Lernpfade, wodurch maßgeschneiderte Bildungserfahrungen für Studierende geschaffen werden (Ng & Fung, 2024; Wen et al., 2024; Zawacki-Richter et al., 2019). Echtzeit-Feedback und automatisierte Anpassungen von Lernstrategien steigern die Lerneffizienz, indem der Lernprozess dynamisch an die Bedürfnisse der Lernenden angepasst wird (Wen et al., 2024). Darüber hinaus führt der Einsatz von KI zu einer Effizienzsteigerung, indem Routineaufgaben automatisiert werden. Automatisierte Bewertungs- und Feedbackprozesse reduzieren den Arbeitsaufwand für Hochschullehrende und verbessern gleichzeitig die Lernleistung der Studierenden (Larrondo et al., 2024; Stamper et al., 2024). Abimbola et al. (2024) betonen, dass Hochschullehrende dadurch mehr Zeit für pädagogische Kernaufgaben gewinnen. Innovative Lehrmethoden mithilfe von Chatbots und LLMs fördern zudem das Lerninteresse und bieten personalisierte Unterstützung (Deng & Yu, 2023; Fryer et al., 2019; Winkler & Söllner, 2018). Diese Technologien erhöhen die Interaktion zwischen Studierenden und Lernmaterialien und unterstützen die Entwicklung akademischer Kompetenzen wie kritisches Denken und Quellenbewertung (Ali et al., 2024; Duran, 2024). Ein Beispiel hierfür ist der KI-gestützte Chatbot ‚Jill Watson‘ an der Georgia Tech University, der Studierende bei häufig gestellten Fragen unterstützt und Lehrkräfte entlastet (Taneja et al., 2024).

Den Potenzialen stehen jedoch erhebliche Herausforderungen gegenüber. Die Wirksamkeit von KI-Systemen hängt stark von der Qualität und der Verfügbarkeit hochwertiger Datensätze ab.

Fehlerhafte oder unvollständige Daten können die Leistungsfähigkeit beeinträchtigen und das Vertrauen in KI-Systeme mindern (Budach et al., 2022). Ressourceneinschränkungen und infrastrukturelle Unterschiede zwischen Institutionen führen außerdem zu Ungleichheiten beim Zugang zu leistungsfähiger Hardware, was bestehende Bildungsungleichheiten verstärken kann (Li, 2023). Zudem wirft der hohe Energieverbrauch moderner KI-Modelle Fragen hinsichtlich der ökologischen Nachhaltigkeit auf (Gowda et al., 2023; Strubell et al., 2019). Darüber hinaus kann der Einsatz von KI zu einem Verlust menschlicher Interaktionen führen, was die Bildungsqualität beeinträchtigen könnte (Duran, 2024).

Die ethischen Implikationen, insbesondere in Bezug auf Datenschutz, sind von zentraler Bedeutung. Die Einhaltung gesetzlicher Vorgaben wie der DSGVO ist essenziell, insbesondere bei der Verarbeitung sensibler personenbezogener Daten (Drachler & Grell, 2016; Bai et al., 2024). Datenschutzverletzungen können das Vertrauen in KI-Systeme erheblich mindern (Slade & Prinsloo, 2013). Algorithmische Verzerrungen (Bias) können zu Ungerechtigkeiten führen und die Chancengleichheit beeinträchtigen (Baker & Hawn, 2022; Omughelli et al., 2024). Transparenz und Nachvollziehbarkeit von KI-Entscheidungen sind daher notwendig, um der ‚Black-Box‘-Problematik entgegenzuwirken und eine gerechte Behandlung aller Studierenden zu gewährleisten (Chaudhary, 2024; Chinta et al., 2024).

Um die Forschungsfrage zu beantworten, wie KI-Systeme, insbesondere LLMs, im akademischen und hochschulischen Kontext konzipiert und implementiert werden sollen, um sowohl Lernprozesse zu verbessern und technische Effizienz zu gewährleisten als auch ethische Standards wie Datenschutz und Fairness einzuhalten, konnten auf Basis aktueller Forschungen drei Aspekte für eine qualitätsvolle Nutzung von KI in der Hochschule herausgearbeitet werden: (1) die Notwendigkeit der Implementierung hochschuleigener KI-Systeme (Modran et al., 2024), (2) aktuelle, zyklisch neu ausgearbeitete Leitlinien (Leitgeb et al., 2024) sowie (3) qualitätsvolle Professionalisierungsmaßnahmen (Umanets et al., 2024; Brandão et al., 2024).

Ad 1) Die Implementierung hochschuleigener KI-Systeme ermöglicht mehrere entscheidende Vorteile:

1. Kontrolle über Daten und Datenschutzkonformität

Durch hochschuleigene KI-Systeme behalten Hochschulen die volle Kontrolle über ihre Daten und können die Einhaltung von Datenschutzrichtlinien wie der DSGVO sicherstellen (Drachler & Grell, 2016). Dies fördert das Vertrauen der Nutzer:innen und ermöglicht den Einsatz der Daten für wissenschaftliche Zwecke innerhalb der Institution. Bai et al. (2024) betonen, dass die Einhaltung gesetzlicher Vorgaben und die Implementierung von Datenschutzmechanismen angesichts neuer KI-Technologien essenziell sind.

2. Anpassung an spezifische Bedürfnisse

Diese KI-Systeme können gezielt auf die spezifischen Anforderungen der Institution zugeschnitten werden, was die Effektivität und die Akzeptanz der Technologie erhöht. Die Anpassung ermöglicht eine bessere Integration in bestehende Lehr- und Lernprozesse und berücksichtigt die individuellen pädagogischen sowie organisatorischen Ziele der Hochschule (Umanets et al., 2024).

3. Bias – systematische Verzerrungen in den Entscheidungen oder Ergebnissen von KI-Systemen

Durch die Nutzung von hochschuleigenen KI-Systemen können Hochschulen gewährleisten, dass die Interaktion mit dem KI-System transparent und frei von diskriminierenden Verzerrungen ist (Holmes et al., 2019). Die Implementierung von erklärbaren KI-Modellen (XAI) erleichtert die Nachvollziehbarkeit von Entscheidungen und fördert die Fairness im Bildungsprozess (Chaudhary, 2024). Omughelli et al. (2024) unterstreichen die Notwendigkeit, Bias in KI-Modellen zu identifizieren und zu minimieren, um qualitätsvolle Ergebnisse zu gewährleisten.

4. Unabhängigkeit und Nachhaltigkeit

Hochschuleigene KI-Systeme reduzieren die Abhängigkeit von externen Anbietern und ermöglichen eine nachhaltige und autonome Entwicklung der KI-Infrastruktur. Dies unterstützt langfristig die Selbstbestimmung der Institution und fördert die Anpassungsfähigkeit an zukünftige technologische Entwicklungen (Li, 2023).

Durch die Umsetzung dieser Maßnahmen können Hochschulen sowohl die technischen Potenziale von KI-Systemen vollständig ausschöpfen als auch die notwendigen ethischen Standards einhalten. Die Implementierung hochschuleigener KI-Systeme stellt somit eine effektive Strategie dar, um technische Effizienz mit ethischer Verantwortung zu vereinen und den Bildungsprozess nachhaltig zu verbessern.

Ad 2) Die Erarbeitung von Leitlinien soll Transparenz und Konsistenz gewährleisten und die akademische Integrität wahren:

Neben der Entwicklung institutionseigener KI-Systeme ist die Etablierung klarer Leitlinien entscheidend für den erfolgreichen und verantwortungsvollen Einsatz von KI in der Hochschulbildung. Leitlinien bieten einen strukturierten Rahmen, der sicherstellt, dass der Einsatz von KI-Technologien im Einklang mit ethischen Grundsätzen und rechtlichen Vorgaben steht. Sie definieren Vorgaben für den Umgang mit Daten, den Einsatz von KI-Tools im Unterricht sowie die Bewertung von KI-generierten Arbeiten, wodurch Transparenz und Konsistenz in der Anwendung gewährleistet werden (Leitgeb et al., 2024). Diese Leitlinien sollten ob der rasanten Entwicklung in diesem Themenbereich zyklisch an diese angepasst werden.

Ad 3) Professionalisierungsmaßnahmen sind unerlässlich:

Professionalisierungsmaßnahmen wie gezielte Schulungen und Weiterbildungsprogramme für Lehrkräfte sind unerlässlich, um das notwendige Wissen und die Kompetenzen für den effektiven Einsatz von KI-Technologien zu vermitteln. Diese Programme sollten nicht nur technische Fähigkeiten abdecken, sondern auch ethische Fragestellungen, Datenschutzbestimmungen und didaktische Methoden zur Integration von KI in den Unterricht umfassen (Umanets et al., 2024; Brandão et al., 2024). Durch kontinuierliche Weiterbildung können Lehrkräfte befähigt werden, KI-Tools verantwortungsvoll und effektiv zu nutzen, wodurch die Qualität der Lehre verbessert und gleichzeitig ethische Standards gewahrt bleiben.

Zusammenfassung

In der vorliegenden Studie wurde untersucht, wie KI und insbesondere LLMs im Bildungsbereich verantwortungsvoll implementiert werden können, um technische Potenziale auszuschöpfen und gleichzeitig ethische Standards einzuhalten. Durch eine systematische Literaturrecherche wurden aktuelle Forschungsergebnisse zu Potenzialen, Herausforderungen und Anforderungen von KI-Systemen in der Hochschulbildung analysiert.

Die aktuelle Literatur zeigt, dass die erfolgreiche Konzeption und Implementierung von KI-Systemen, insbesondere von LLMs, im akademischen Kontext auf drei wesentlichen Säulen beruht: der Entwicklung hochschuleigener KI-Systeme, der Etablierung klarer Leitlinien sowie der Durchführung qualitativvoller Professionalisierungsmaßnahmen. Erstens ermöglicht die Entwicklung eigener KI-Systeme den Hochschulen die volle Kontrolle über ihre Daten und die Anpassung der Technologien an spezifische institutionelle Bedürfnisse, was sowohl die technische Effizienz als auch die Einhaltung ethischer Standards unterstützt (Modran et al., 2024; Drachler & Greller, 2016). Zweitens sorgen regelmäßig aktualisierte Leitlinien für Transparenz und Konsistenz im Umgang mit KI. Sie wahren die akademische Integrität und stellen die Einhaltung rechtlicher Vorgaben sicher (Leitgeb et al., 2024). Drittens sind umfassende Professionalisierungsmaßnahmen unerlässlich, um Lehrkräfte mit den notwendigen technischen und ethischen Kompetenzen auszustatten, sodass sie KI-Tools verantwortungsvoll in den Unterricht integrieren können (Umanets et al., 2024; Brandão et al., 2024). Durch die gezielte Kombination dieser drei Aspekte können Hochschulen die Potenziale von KI-Technologien effektiv nutzen, Lernprozesse optimieren und gleichzeitig hohe ethische Standards wahren.

Obwohl diese Studie wesentliche Einblicke in die verantwortungsvolle Implementierung von KI und LLMs im Bildungsbereich bietet, gibt es mehrere Limitationen, die berücksichtigt werden müssen:

1. **Eingeschränkte Datenbasis:** Die systematische Literaturrecherche basiert auf einer Auswahl verfügbarer Studien und Publikationen, die zwischen 2014 und 2024 veröffentlicht wurden. Diese zeitliche Begrenzung könnte dazu führen, dass relevante neuere Entwicklungen, die nach Abschluss der Recherche publiziert wurden, nicht berücksichtigt sind.
2. **Theoretische Fokussierung:** Die Studie bezieht sich überwiegend auf theoretische und konzeptionelle Ansätze zur Implementierung von KI im Bildungswesen. Empirische Studien, die reale Implementierungen von KI-Systemen in Bildungseinrichtungen bewerten, sind vergleichsweise weniger vertreten. Dies schränkt die Übertragbarkeit der Empfehlungen auf praktische Kontexte ein.
3. **Heterogenität der Quellen:** Die in der systematischen Literaturrecherche berücksichtigten Studien variieren hinsichtlich ihrer Methodik und Anwendungsbereiche, was eine Limitation darstellt, da dies die Vergleichbarkeit der Ergebnisse erschwert und die Generalisierbarkeit der Schlussfolgerungen einschränkt.
4. **Fehlende Langzeitstudien:** Die meisten analysierten Studien befassen sich mit kurzfristigen Ergebnissen oder Pilotprojekten zur Einführung von KI in Bildungseinrichtungen. Langzeitstudien, die die langfristigen Auswirkungen der Implementierung von KI auf das Bildungssystem und die Lernenden untersuchen, fehlen weitgehend.

5. **Unklare Kosten-Nutzen-Analyse:** Obwohl die Studie Handlungsempfehlungen für die Investition in technische Infrastruktur und Schulungen gibt, wird der finanzielle Aufwand, insbesondere für kleinere oder weniger finanzstarke Bildungseinrichtungen, nicht ausreichend analysiert. Die wirtschaftliche Machbarkeit sowie die Kosten-Nutzen-Verhältnisse werden nur am Rande erwähnt.
6. **Ethische Implikationen:** Während ethische Herausforderungen wie algorithmische Verzerrungen und Datenschutz angesprochen werden, bleibt die Analyse dieser Themen teilweise auf einem allgemeinen Niveau. Konkrete, praxisorientierte Lösungsansätze, insbesondere hinsichtlich der Einhaltung komplexer gesetzlicher Vorschriften wie der DSGVO, werden nicht detailliert untersucht.

Die genannten Limitationen verdeutlichen, dass weitere Forschung, vor allem empirische Studien und Langzeituntersuchungen, erforderlich ist, um die Auswirkungen von KI und LLMs im Bildungsbereich umfassender zu verstehen und praxisnähere Handlungsempfehlungen zu entwickeln.

Jedoch liefert die vorliegende Studie durchaus praxisnahe Leitlinien zur verantwortungsvollen Integration von KI-Technologien im Bildungsbereich, mit einem besonderen Fokus auf LLMs. Sie zeigt, dass Hochschulen durch gezielte Maßnahmen sowohl technische Potenziale ausschöpfen als auch ethische und datenschutzrechtliche Standards wahren können. Zu den zentralen Handlungsempfehlungen zählen Investitionen in die technische Infrastruktur, die Förderung interdisziplinärer Zusammenarbeit, die Schulung von Lehrkräften in technischen und ethischen Kompetenzen sowie die Etablierung von Leitlinien. Diese Schritte ermöglichen es, KI-Systeme erfolgreich in den Lehrbetrieb zu integrieren, die Qualität der Bildung zu verbessern und gleichzeitig hohe ethische und rechtliche Anforderungen zu erfüllen.

Literaturverzeichnis

A

- Abimbola, C., C. A., Chisom, O. N., & Adeniyi, I. S. (2024). Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Reviews*, 10(2), 006-013. <https://doi.org/10.30574/msarr.2024.10.2.0039>
- Akpan, I. J., Kobara, Y. M., Owolabi, J., Akpan, A. A., & Offodile, O. F. (2024). Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *International Transactions in Operational Research*. <https://doi.org/10.1111/itor.13522>
- Al Darayseh, A. (2023). Acceptance of artificial intelligence in teaching science: Science teachers' perspective. *Computers and Education: Artificial Intelligence*, 4, 100132. <https://doi.org/10.1016/j.caeai.2023.100132>
- Ali, M. H., Kainat, M., Maqsood, M., Fernandez, S. F., Ali, S., & Ahmad, M. (2024). Chatbots in Academic Literacy Exploring AI in Education. In Ş. Demir & M. Demir (Eds.), *Enhancing Higher Education and Research With OpenAI Models* (pp. 37-59). IGI Global. <https://doi.org/10.4018/979-8-3693-1666-5.ch002>
- Atlas, S. (2023). *ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI*. https://digitalcommons.uri.edu/cba_facpubs/548

B

- Baker, R. S., & Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 31(4), 595–626. <https://doi.org/10.1007/s40593-021-00285-9>
- Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability*, 15, 12983. <https://doi.org/10.3390/su151712983>
- Bai, J. Y. H., Zawacki-Richter, O., & Muskens, W. (2024). RE-EXAMINING THE FUTURE PROSPECTS OF ARTIFICIAL INTELLIGENCE IN EDUCATION IN LIGHT OF THE GDPR AND ChatGPT. *Turkish Online Journal of Distance Education*, 25(1), 20-32. <https://doi.org/10.17718/tojde.1248901>
- Bañeres, D., Rodríguez-González, M. E., Guerrero-Roldán, A. E., & Cortadas, P. (2023). An early warning system to identify and intervene online dropout learners. *International Journal of Educational Technology in Higher Education*, 20(1), 1-25. <http://dx.doi.org/10.1186/s41239-022-00371-5>
- Bhattacharya, A., Stumpf, S., & Verbert, K. (2024). Representation Debiasing of Generated Data Involving Domain Experts. *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*. <https://doi.org/10.48550/arXiv.2407.09485>
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (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(4). <https://doi.org/10.1007/s10639-022-10925-9>

- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 149–159). <http://dx.doi.org/10.48550/arXiv.1712.03586>
- Brandão, A., Pedro, L., & Zagalo, N. (2024). Teacher professional development for a future with generative artificial intelligence – An integrative literature review. *Digital Education Review*, 45(June), 151-157. <https://doi.org/10.1344/der.2024.45.151-157>
- Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., Harmouch, H., & Naumann, F. (2022). The effects of data quality on machine learning performance. <https://doi.org/10.48550/arxiv.2207.14529>

C

- Chaudhary, G. (2024). Unveiling the black box: Bringing algorithmic transparency to AI. *Masaryk University Journal of Law and Technology*, 18(1). 93-122. <https://doi.org/10.5817/mujlt2024-1-4>
- Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to Re-visit Existing Student's Performance Evaluation Approach in Higher Education Sector in a New Era of ChatGPT—A Case Study. *Cogent Education*, 10(1), 2210461. <http://dx.doi.org/10.1080/2331186X.2023.2210461>
- Chinta, S. V., Wang, Z., Yin, Z., Hoang, N., Gonzalez, M., Le Quy, T., & Zhang, W. (2024). FairAIED: Navigating fairness, bias, and ethics in educational AI applications. <https://doi.org/10.48550/arxiv.2407.18745>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. <http://dx.doi.org/10.1109/ACCESS.2020.2988510>
- Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22–42. <https://doi.org/10.14742/ajet.7526>
- Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(2), 02. <https://ro.uow.edu.au/cgi/viewcontent.cgi?article=3180&context=jutlp>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20, 22. <https://doi.org/10.1186/s41239-023-00392-8>

D

- Da Cruz, R. C., Juliano, R. C., Monteiro Souza, F. C., & Correa Souza, A. C. (2023, Mai). A Score approach to identify the risk of students dropout: An experiment with Information Systems Course. In *Proceedings of the XIX Brazilian Symposium on Information Systems* (S. 120-127). <https://doi.org/10.1145/3592813.3592896>
- D'ascenzo, F., Rocchi, A., Iandolo, F., & Vito, P. (2024). Evolutionary impacts of artificial intelligence in Healthcare Managerial Literature. *Sustainability Futures*, 7, 100198. <https://doi.org/10.1016/j.sftr.2024.100198>
- Daun, M., & Brings, J. (2023). How ChatGPT will change software engineering education. In *Proceedings of the 2023 Conference on Innovation and Technology in*

Computer Science Education V. 1 (pp. 110-116).

<http://dx.doi.org/10.1145/3587102.3588815>

- Dave, T., Athaluri, S. A., & Singh, S. (2023). ChatGPT in medicine: An overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in Artificial Intelligence*, 6, 1169595. <https://doi.org/10.3389/frai.2023.1169595>
- Deng, X., and Yu, Z. (2023). A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education. *Sustainability* 15, no. 4: 2940. <https://doi.org/10.3390/su15042940>
- Drachler, H., & Greller, W. (2016). Privacy and analytics: It's a delicate issue. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 89–98). <https://doi.org/10.1145/2883851.2883893>
- Duran, V. (2024). Analyzing teacher candidates' arguments on AI integration in education via different chatbots. <https://doi.org/10.1344/der.2024.45.68-83>

E

- Estévez-Ayres, I., Callejo, P., Hombrados-Herrera, M. Á., Alario-Hoyos, C., & Delgado Kloos, C. (2024). Evaluation of LLM tools for feedback generation in a course on concurrent programming. <https://doi.org/10.1007/s40593-024-00406-0>

F

- Farhat, F., Silva, E. S., Hassani, H., Madsen, D., Sohail, S. S., Himeur, Y., Alam, M. A., & Zafar, A. (2023). The scholarly footprint of ChatGPT: A bibliometric analysis of the early outbreak phase. *Frontiers in Artificial Intelligence*, 6, 1270749. <https://doi.org/10.3389/frai.2023.1270749>
- Farooqi, M. T. K., Amanat, I., & Awan, S. M. (2024). Ethical Considerations and Challenges in the Integration of Artificial Intelligence in Education: A Systematic Review. *Journal of Excellence in Management Sciences*, 3(4), 35–50. <https://doi.org/10.69565/jems.v3i4.314>
- Floridi, L., Cowls, J. (2021). A Unified Framework of Five Principles for AI in Society. In: Floridi, L. (eds) *Ethics, Governance, and Policies in Artificial Intelligence*. Philosophical Studies Series, vol 144. Springer, Cham. https://doi.org/10.1007/978-3-030-81907-1_2
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289. <https://doi.org/10.1016/j.chb.2018.12.023>

G

- Gal, E., Kaplan-Rakowski, R., Amzalag, M., Shaked, N., Zaguri, Y., Kohen-Vacs, D., Barak-Medina, E., & Kurtz, G. (2024). Strategies for Integrating Generative AI into Higher Education: Navigating Challenges and Leveraging Opportunities. *Education Sciences*, 14, 503. <https://doi.org/10.3390/educsci14050503>
- Gonzalez-Nucamendi, A., Noguez, J., Neri, L., Robledo-Rella, V., & García-Castelán, R. M. G. (2023). Predictive Analytics Study to Determine Undergraduate Students at Risk of Dropout. *Frontiers in Education*, 8, 1244686. <https://doi.org/10.3389/feduc.2023.1244686>

- Go, M. B., Junior, R. A. G., Velos, S. P., Dayupay, J. P., Cababat, F. G., Baird, J. C. C., & Quiñanola, H. (2023). A data mining approach to classifying e-learning satisfaction of higher education students: a Philippine case. *International Journal of Innovation and Learning*, 33(3), 314-329. <https://EconPers.repec.org/RePEc:ids:ijilea:v:33:y:2023:i:3:p:314-329>
- Gowda, S.N., Hao, X., Li, G., Sevilla-Lara, L., & Gowda, S.N. (2023). Watt For What: Rethinking Deep Learning's Energy-Performance Relationship. <https://doi.org/10.48550/arXiv.2310.06522>
- Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*, (2nd ed.). Los Angeles: SAGE. https://uk.sagepub.com/sites/default/files/upm-assets/81596_book_item_81596.pdf
- Goel, P. K. (2024). AI for Energy Efficiency and Conservation. In B. Riswandi, B. Singh, C. Kaunert, & K. Vig (Eds.), *AI Applications for Clean Energy and Sustainability* (pp. 32-49). IGI Global. <https://doi.org/10.4018/979-8-3693-6567-0.ch003>

H

- Hidalgo, I., Fernández-de-Vega, F., Ceberio, J., Garnica, Ó., Velasco, J. M., Cortés, J. C., Villanueva, R., & Díaz, J. (2023, December 7). Sustainable Artificial Intelligence Systems: An Energy Efficiency Approach. *Universidad Complutense de Madrid*. <https://doi.org/10.36227/techrxiv.24610899.v1>
- Hinojo-Lucena, F. J., Aznar-Díaz, I., Cáceres-Reche, M. P., & Romero-Rodríguez, J. M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, 9(1), 51. <https://doi.org/10.3390/educsci9010051>
- Holmes, W. (2019). Artificial Intelligence in Education. In Tatnall, A. (Ed.), *Encyclopedia of Education and Information Technologies*. Springer, Cham. https://doi.org/10.1007/978-3-319-60013-0_107-1
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and Implications for Teaching and Learning*. Center for Curriculum Redesign. https://www.researchgate.net/publication/332180327_Artificial_Intelligence_in_Education_Promise_and_Implications_for_Teaching_and_Learning
- Hung, J., & Chen, J. (2023). The Benefits, Risks and Regulation of Using ChatGPT in Chinese Academia: A Content Analysis. *Social Sciences*, 12, 380. <https://doi.org/10.3390/socsci12070380>

I

- IMS Global Learning Consortium. (2020). Learning tools interoperability. Abgerufen von <https://www.imsglobal.org/activity/learning-tools-interoperability>
- Ioku, T., Kondo, S., & Watanabe, Y. (2024). Acceptance of generative ai in higher education: a latent profile analysis of policy guidelines. <https://doi.org/10.21203/rs.3.rs-4515787/v1>

J

K

- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., Papin, K., & Kurtz, G. (2024). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, 34, 313–338. [https://www.academia.edu/106455593/Generative AI and Teachers Perspectives on Its Implementation in Education](https://www.academia.edu/106455593/Generative_AI_and_Teachers_Perspectives_on_Its_Implementation_in_Education)
- Kasneci, E., Sessler, K., Kasneci, G., & Seidel, T. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- KI@PPHB. (2024a). Interaktion mit KI@PPHB zwischen 28.06-30.06.2024 (GPT4o). Input: [Kontrolliere nachfolgende Liste von Publikationen und identifiziere doppelte Studien (schrittweises Vorgehen – jeweils 100 Titeln)]. Output: [Identifikation doppelter Titel]. <https://ai.ph-burgenland.at>
- KI@PPHB. (2024b). Interaktion mit KI@PPHB zwischen 01.07-04.08.2024 (GPT4o). Input: [Fasse mir die Abstracts zusammen und füge mir den Titel hinzu. Maximal 2 Sätze. Wichtig sind die Stichprobe die Methodik und das Ziel der Studie]. Output: [Zusammenfassung der Abstracts mit Titeln]. <https://ai.ph-burgenland.at>
- KI@PPHB. (2024c). Interaktion mit KI@PPHB zwischen 01.07-04.08.2024 (GPT4o). Input: [Untersuche nachfolgende Abstracts, Identifiziere Studien, Artikeln, Konferenzbeiträge, die sich KI-Anwendungen in der Bildung beschäftigen (Zahlreiche Iterationen waren nötig)]. Output: [Identifikation der Titeln]. <https://ai.ph-burgenland.at>

L

- Larrondo, P., Frank, B., & Ortiz, J. M. (2024). *Work-in-progress: Fine-tuning large language models for automated feedback in complex engineering problem-solving*. The Future Of Engineering Education – Annual Conference and Exposition 2024.
- Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine*, 388, 1233–1239. <https://doi.org/10.1056/nejmsr2214184>
- Leitgeb, T., Maitz, K., Sitter, G., Matischek-Jauk, M., Möblacher, C., Knaus, M., Gabriel, H., & Meller, S. (2024). *KI-Leitlinien für den PH-Verbund Süd-Ost – Leitlinien für die Nutzung von Künstlicher Intelligenz in der Hochschule*. University College of Teacher Education Burgenland, Eisenstadt. <https://doi.org/10.5281/zenodo.13693466>
- Li, H. (2023). AI in Education: Bridging the Divide or Widening the Gap? Exploring Equity, Opportunities, and Challenges in the Digital Age. *Advances in Education, Humanities and Social Science Research*. <https://doi.org/10.56028/aehtsr.8.1.355.2023>
- Lipton, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM*, 61(10), 36–43. <https://doi.org/10.1145/3233231>

M

- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12. Aufl.). Beltz.
- Mayring, P. (2019). Qualitative Inhaltsanalyse. Abgrenzungen, Spielarten, Weiterentwicklungen [30 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), Art. 16. <http://dx.doi.org/10.17169/fqs-20.3.3343>

- McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). *Generative Artificial Intelligence in Higher Education: Evidence from an Analysis of Institutional Policies and Guidelines*. <https://doi.org/10.48550/arXiv.2402.01659>
- Misino, E., Calegari, R., Lombardi, M., & Milano, M. (2024). Ensuring fairness stability for disentangling social inequality in access to education: The FAIRDAS general method. *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 7412–7420. <https://doi.org/10.24963/ijcai.2024/820>
- Modran, H., Bogdan, I. C., Ursuțiu, D., Samoila, C., & Modran, P. L. (2024). LLM Intelligent Agent Tutoring in Higher Education Courses using a RAG Approach. *Preprints*. <https://doi.org/10.20944/preprints202407.0519.v1>
- Mendiratta, A., Singh, S., Yadav, S. S., & Mahajan, A. (2023). Bibliometric and Topic Modeling Analysis of Corporate Social Irresponsibility. *Global Journal of Flexible Systems Management*, 2022, 1–21. <https://doi.org/10.1007/s40171-023-00343-2>

N

- Natukunda, A., & Muchene, L. K. (2023). Unsupervised title and abstract screening for systematic review: A retrospective case-study using topic modelling methodology. *Systematic Reviews*, 12, 1–16. <https://doi.org/10.1186/s13643-022-02163-4>
- Ng, C., & Fung, Y. (2024). *Educational Personalized Learning Path Planning with Large Language Models*. <https://arxiv.org/abs/2407.11773>
- Niu, K., Jia, B., Zhou, Y., & Lu, G. (2022). A hybrid model for predicting academic performance of engineering undergraduates. *Int. J. Model. Simul. Sci. Comput.*, 14, 2350030:1–2350030:19. <http://dx.doi.org/10.1142/S1793962323500307>

O

- Oliński, M., Krukowski, K., & Sיעiński, K. (2024). Bibliometric Overview of ChatGPT: New Perspectives in Social Sciences. *Publications*, 12, 9. <https://doi.org/10.3390/publications12010009>
- Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893–7925. <https://doi.org/10.1007/s10639-022-10925-9>
- Ogunleye, B., Zakariyyah, Kl., Ajao, O., Olayinka, O., & Sharma, H. (2024). A Systematic Review of Generative AI for Teaching and Learning Practice. *Education Sciences* 14, no. 6: 636. <https://doi.org/10.3390/educsci14060636>
- Omughelli, D., Gordon, N., & Al Jaber, T. (2024). Fairness, Bias, and Ethics in AI: Exploring the Factors Affecting Student Performance. *Journal of Intelligent Communication*, 4(1), 100–110. <https://doi.org/10.54963/jic.v4i1.306>
- OpenAI. (2023). *GPT-4 Technical Report*. <https://doi.org/10.48550/arxiv.2303.08774>

P

- Pagano, T., Loureiro, R. B., Lisboa, F. V. N., Peixoto, R. M., Guimarães, G. A. S., Cruz, G. O. R., Araujo, M. M., Santos, L. L., Cruz, M. A. S., Oliveira, E. L. S., Winkler, I., & Nascimento, E. G. S. (2023). Bias and unfairness in machine learning models: A systematic review. *Big Data Cogn. Comput.*, 7 (15). <https://doi.org/10.3390/bdcc7010015>

- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journal of Mass Communication Education*, 78, 84–93. <https://doi.org/10.1177/10776958221149577>
- Pannu, J., & Boosalis, C. (2023). A Use-Case for Implementing ChatGPT to Augment Teaching an Introductory Statistics Course. In *Proceedings of the Future Technologies Conference* (S. 196-203). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47454-5_15
- Pishtari, G., Wagner, M., & Ley, T. (2024). Bericht für Arbeitspaket 3 (Preprint) Ein Forschungsüberblick über den Einsatz von Künstlicher Intelligenz für das Lehren und Lernen in der Hochschulbildung. In Gerhard Brandhofer, Ortrun Gröbinger, Tanja Jadin, Michael Raunig & Julia Schindler (Eds.), *Von KI lernen, mit KI lehren: Die Zukunft der Hochschulbildung*. FNMA, Lustenau.
- PRISMA. (2020). Preferred Reporting Items for Systematic Reviews and Meta-Analyses. <http://www.prisma-statement.org/>

R

- Romanelli, J. P., Gonçalves, M. C. P., de Abreu Pestana, L. F., Soares, J. A. H., Boschi, R. S., Andrade, D. F. (2021). Four challenges when conducting bibliometric reviews and how to deal with them. *Environmental Science and Pollution Research*, 28, 60448–60458. <https://doi.org/10.1007/s11356-021-16420-x>
- Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta Llovet, J., Gonzalez-Rodriguez, M., et al. (2023). Empowering Education with Generative Artificial Intelligence Tools: Approach with an Instructional Design Matrix. *Sustainability*, 15, 11524. <https://doi.org/10.3390/su151511524>

S

- Samek, W., Wiegand, T., & Müller, K.-R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *ITU Journal: ICT Discoveries, Special Issue 1*, 39–48. <http://dx.doi.org/10.48550/arXiv.1708.08296>
- Sanasintani, S. (2023). Revitalizing The Higher Education Curriculum Through An Artificial Intelligence Approach: An Overview. *Journal of Social Science Utilizing Technology*. <http://dx.doi.org/10.55849/jssut.v1i4.670>
- Selwyn, N. (2019). *Should robots replace teachers? AI and the Future of Education*. (1st ed.) Polity Press. <https://www.wiley.com/en-gb/Should+Robots+Replace+Teachers%3F%3A+AI+and+the+Future+of+Education-p-9781509528967>
- Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, 57(10), 1510- 1529. <https://doi.org/10.1177/0002764213479366>
- Smart Sparrow. (2024). Adaptive learning platform. Abgerufen von <https://www.smartsparrow.com>
- Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning. *Learning and Instruction*, 83, 101620. <https://doi.org/10.1016/j.learninstruc.2022.101620>
- Singh, M., Sharma, D., Ma, A., & Goldfine, N. (2024). *Towards More Accurate Prediction of Human Empathy and Emotion in Text and Multi-turn Conversations by*

Combining Advanced NLP, Transformers-based Networks, and Linguistic Methodologies. <https://arxiv.org/abs/2407.18496>

- Singhal, S. (2024). Data Privacy, Compliance, and Security Including AI ML: Healthcare. In P. Whig, S. Sharma, S. Sharma, A. Jain, & N. Yathiraju (Eds.), *Practical Applications of Data Processing, Algorithms, and Modeling* (pp. 111-126). IGI Global. <https://doi.org/10.4018/979-8-3693-2909-2.ch009>
- Stamper, J., Xiao, R., & Hou, X. (2024). Enhancing LLM-based feedback: Insights from intelligent tutoring systems and the learning sciences. <https://doi.org/10.48550/arxiv.2405.04645>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6, 1–11. <http://dx.doi.org/10.37074/jalt.2023.6.1.17>
- Suresh, S., & Misra, S. M. (2024). Large Language Models in Pediatric Education: Current Uses and Future Potential. *Pediatrics*, 154(3), e2023064683. <https://doi.org/10.1542/peds.2023-064683>

T

- Tanisha, J., Rajesh, P., Singh, R., Adhip, K., Stuti, K., & Ajitha, D. (2024). Privacy and data protection challenges in Industry 4.0: An AI-driven perspective. *World Journal of Advanced Engineering Technology and Sciences*. <https://doi.org/10.30574/wjaets.2024.12.2.0287>
- Taneja, K., Maiti, P., Kakar, S., Guruprasad, P., Rao, S., Goel, A.K. (2024). Jill Watson: A Virtual Teaching Assistant Powered by ChatGPT. In: Olney, A.M., Chounta, IA., Liu, Z., Santos, O.C., Bittencourt, I.I. (eds) *Artificial Intelligence in Education*. AIED 2024. Lecture Notes in Computer Science(), vol 14829. Springer, Cham. https://doi.org/10.1007/978-3-031-64302-6_23
- Thurzo, A., Strunga, M., Urban, R., Surovková, J., Afrashtehfar, K. I. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum update. *Education Sciences*, 13, 150. <http://dx.doi.org/10.3390/educsci13020150>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Tulang, R., & Chen, Y. (2023). AI in Education: Bridging the Divide or Widening the Gap? Exploring Equity, Opportunities, and Challenges in the Digital Age. *Advances in Education, Humanities and Social Science Research*, 8(1), 355–370. <https://doi.org/10.56028/aehtsr.8.1.355.2023>

U

- Umanets, V., Shakhina, I., & Rozputnia, B. (2024). *Modern Information Technologies and Innovation Methodologies of Education in Professional Training Methodology Theory Experience Problems*, 72, 162-169. <https://doi.org/10.31652/2412-1142-2024-72-162-170>

V

- Valentin, S., Fu, J., Detommaso, G., Xu, S., Zappella, G., & Wang, B. (2024). Cost-Effective Hallucination Detection for LLMs. <https://arxiv.org/abs/2407.21424v2>
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84, 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Van Eck, N. J., & Waltman, L. (2017). Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics*, 111, 1053–1070. <http://dx.doi.org/10.1007/s11192-017-2300-7>
- Villarreal-Torres, H., Ángeles-Morales, J., Marín-Rodríguez, W., Andrade-Girón, D., Cano-Mejía, J., Mejía-Murillo, C., & Palomino-Márquez, M. (2023). Classification model for student dropouts using machine learning: A case study. *EAI Endorsed Transactions on Scalable Information Systems*. <http://dx.doi.org/10.4108/eetsis.vi.3455>
- Vagelas, I., & Leontopoulos, S. (2024). A Bibliometric Analysis and a Citation Mapping Process for the Role of Soil Recycled Organic Matter and Microbe Interaction due to Climate Change Using Scopus Database. *AgriEngineering*, 5, 581–610. <http://dx.doi.org/10.3390/agriengineering5010037>
- Villarreal-Torres, H., Ángeles-Morales, J., Marín-Rodríguez, W., Andrade-Girón, D., Cano-Mejía, J., Mejía-Murillo, C., & Palomino-Márquez, M. (2023). Classification model for student dropouts using machine learning: A case study. *EAI Endorsed Transactions on Scalable Information Systems*. <http://dx.doi.org/10.4108/eetsis.vi.3455>
- Vincent-Lancrin, S. and R. van der Vlies (2020), "Trustworthy artificial intelligence (AI) in education: Promises and challenges", *OECD Education Working Papers*, No. 218, OECD Publishing, Paris, <https://doi.org/10.1787/a6c90fa9-en>.
- von, Eschenbach., J, Warren. (2021). Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology*, 1-16 <https://link.springer.com/article/10.1007%2Fs13347-021-00477-0>.

W

- Wang, L., Ma, C., Feng, X. *et al.* A survey on large language model based autonomous agents. *Front. Comput. Sci.* 18, 186345 (2024). <https://doi.org/10.1007/s11704-024-40231-1>
- Wen, Q., Liang, J., Sierra, C., Luckin, R., Tong, R., Liu, Z., Cui, P., & Tang, J. (2024). AI for education (AI4EDU): Advancing personalized education with LLM and adaptive learning. <https://doi.org/10.1145/3637528.3671498>
- Wen, W., Liu, Y., Zhu, Z., & Shi, Y. (2023). A Study on the Learning Early Warning Prediction Based on Homework Habits: Towards Intelligent Sustainable Evaluation for Higher Education. *Sustainability*, 15(5), 4062. <https://doi.org/10.3390/su15054062>
- Wei, L. (2023). Artificial intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1261955. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Proceedings of the International Conference on Information Systems*. <https://doi.org/10.5465/AMBPP.2018.15903abstract>
- Wuerfe, T., & Mayweg-Paus, E. (2023). “Your argumentation is good”, says the AI vs humans–The role of feedback providers and personalised language for feedback effectiveness. *Computers and Education: Artificial Intelligence*, 5, 100189. <https://doi.org/10.1016/j.caeai.2023.100189>

X

Y

- Yang, H., Gao, C., & Shen, H. Z. (2023). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, 1-22.
<https://doi.org/10.1007/s10639-022-10925-9>

Z

- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (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>
- Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., & Hong, C. S. (2023). One small step for generative AI, one giant leap for AGI: A complete survey on ChatGPT in AIGC era. *arXiv, abs/2304.06488*. <https://arxiv.org/abs/2304.06488>
- Zheng, L., Niu, J., Zhong, L., & Gyasi, J.F. (2021). The effectiveness of artificial intelligence on learning achievement and learning perception: A meta-analysis. *Interactive Learning Environments*, 31, 5650 - 5664.
<http://dx.doi.org/10.1080/10494820.2021.2015693>