



Bildung nach Maß: Ein Ansatz zur Adaption von Lerntexten durch generative Künstliche Intelligenz

Daniel Stattkus · Laura Hein · Oliver Thomas

Eingegangen: 7. August 2024 / Angenommen: 8. Januar 2025
© The Author(s) 2025

Zusammenfassung Der vorliegende Artikel thematisiert das von den Vereinten Nationen definierte Ziel einer hochwertigen Bildung durch den Einsatz Generativer Künstlicher Intelligenz zur Adaption von Lerntexten. Durch die Anpassung von Lerntexten an individuelle und gruppenbasierte Interessen soll deren Zugänglichkeit und Relevanz erhöht werden. Zur Umsetzung dieser Adaption wird ein Large Language Model mit angepasstem Prompt verwendet. Die Evaluationsergebnisse zeigen, dass individualisierte Texte im Vergleich zu standardisierten Texten besser angenommen werden, wohingegen Anpassungen auf Gruppenebene weniger effektiv sind. Um zukünftig das Potenzial der Technologie voll auszuschöpfen und eine sorgfältige Abstimmung zwischen den Interessen der Lernenden und den Lerninhalten zu gewährleisten, ist eine enge Zusammenarbeit zwischen Technologie und Didaktik unter Einbeziehung der Lehrkräfte erforderlich.

Schlüsselwörter Generative Künstliche Intelligenz · Bildung · Individualisierung · Differenzierung · Adaption

Daniel Stattkus · Laura Hein · Oliver Thomas
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), DFKI Niedersachsen,
Smart Enterprise Engineering, Hamburger Straße 24, 49084 Osnabrück, Deutschland

Daniel Stattkus
E-Mail: daniel.stattkus@dfki.de

Oliver Thomas
E-Mail: oliver.thomas@uni-osnabrueck.de

✉ Laura Hein · Oliver Thomas
Fachgebiet Informationsmanagement und Wirtschaftsinformatik, Universität Osnabrück, Hamburger
Straße 24, 49084 Osnabrück, Deutschland
E-Mail: laura.hein@dfki.de

Customized Education: An Approach to Adapt Learning Texts through Generative Artificial Intelligence

Abstract This article addresses the United Nations' goal of high-quality education through the use of Generative Artificial Intelligence to adapt learning texts. By tailoring learning materials to individual and group-based interests, their accessibility and relevance are to be enhanced. To implement this adaptation, a Large Language Model with a customized prompt is used. Evaluation results show that personalized texts are more accepted compared to standardized texts, whereas adjustments at the group level are less effective. To fully leverage the potential of this technology in the future and ensure careful alignment between learners' interests and the learning content, close collaboration between technology and educational methods involving teachers is required.

Keywords Generative Artificial Intelligence · Education · Individualization · Differentiation · Adaptation

1 Potenziale der Generativen Künstlichen Intelligenz: Lerntextadaption zur Verbesserung von Lernprozessen

In den von den Vereinten Nationen definierten Nachhaltigkeitszielen wird ein Fokus auf die inklusive, chancengerechte und hochwertige Bildung gelegt. So weisen sie auf die Bedeutung der Bildung in Hinblick auf zukünftige politische, soziale, kulturelle und wirtschaftliche Situationen hin (Vereinte Nationen 2024). Um die Bildungsqualität zu verbessern, konzentriert sich dieser Artikel auf die Adaption von Lerntexten durch den Einsatz von Generativer Künstlicher Intelligenz (GKI), um zielgruppengerecht auf die Bedürfnisse der Lernenden einzugehen. Die Auswahl von Lerntexten als Untersuchungsgegenstand wurde bewusst getroffen, um die Prinzipien der Adaption an einem möglichst einfachen, universell verständlichen und bekannten Lerninhalt zu demonstrieren. GKI hat das Potenzial, den Zugang zu Bildung erheblich zu verbessern, indem sie Lehrmaterialien adaptiert und damit nutzerorientierter und für jeden einzelnen Lernenden zugänglicher gestaltet. Die vorliegende Studie verfolgt das Ziel, Lerntexte auf die individuellen Interessen der Lernenden zuzuschneiden. Zu diesem Zweck wird ein Large Language Model (LLM) verwendet. Dabei werden vier Gruppen von Texten verglichen. Standard-Texte, die nicht adaptiert wurden (SLT), Texte die spezifisch auf die Interessen des Individuums angepasst wurden (ILT), Texte, die für eine Gruppe individualisiert wurden, der das Individuum angehört (GLT), und Texte, die für eine der anderen Gruppen angepasst wurden, der das Individuum nicht angehört (FLT). Die FLT dienen dazu, die Gefahren einer falschen Adaption zu untersuchen. Basierend auf diesen Gruppen stellen wir folgende Forschungsfragen auf:

1. Wie effektiv kann GKI Lerntexte an die Interessen von Lernenden anpassen?
2. Welchen Einfluss hat die Adaption von Lerntexten durch GKI auf Lernende?
3. Wie wirkt sich eine fehlerhafte Adaption auf die Lernenden aus?

Um diese Forschungsfragen zu beantworten, ist dieser Artikel wie folgt aufgebaut: Zunächst wird in Kap. 2 die Rolle von Technologien zur Adaption von Lerninhalten erläutert. Kap. 3 beschreibt die Umsetzung der Adaption mit GKI. In Kap. 4 erläutern wir die durchgeführte Evaluation. Anschließend werden in den Kap. 5 und 6 die wichtigsten Ergebnisse zusammengefasst und diskutiert. Abschließend wird in Kap. 7 ein Fazit gezogen sowie die Limitationen dieser Studie erläutert und ein Ausblick auf zukünftige Forschung gegeben.

2 Die Rolle von Technologien zur Adaption von Lerninhalten

Die Adaption von Lerninhalten durch flexible Lehrpläne und Methoden ist ein in der Didaktik weit verbreitetes Konzept und wird dort häufig als äußere oder innere Differenzierung bezeichnet. Die äußere Differenzierung bezieht sich auf die Unterteilung der Lernenden z.B. nach Schulformen. Bei der inneren Differenzierung hingegen werden die Lehrmethoden, Inhalte oder Lernziele an die unterschiedlichen Lernvoraussetzungen und Bedürfnisse der einzelnen Lernenden angepasst (Klafki und Stöcker 1976; Streber 2015; Tomlinson 2001). Differenzierung durch Adaption in der Bildung ist essenziell für die Kreativitäts- und Kompetenzentwicklung der Lernenden (Li und Jiang 2013). Eine Studie von Karpenko (2019) schlägt vor, dass die Adaption in der Bildung durch den gezielten Einsatz von Technologien wie Big Data und intelligenten Informationssystemen weiter vorangetrieben werden kann. Indem individuelle Lernprofile analysiert werden, können Lehrpläne und Lehrmethoden dynamisch angepasst werden. Darüber hinaus ermöglichen KI-gestützte Wissensbewertungsmethoden, wie sie von Minn (2022) untersucht wurden, durch formative und summative Assessments eine präzisere Anpassung von Lerninhalten an die individuellen Bedürfnisse der Lernenden und schaffen so eine Grundlage für skalierbare adaptive Lernsysteme. Weitere Forschungsarbeiten betonen, dass durch den Einsatz von Bildungstechnologien den Lehrkräften ermöglicht wird, den Fortschritt der Lernenden zu überwachen und den Unterricht entsprechend anzupassen. Darüber hinaus verbessern interaktive Technologien wie Learning-Management-Systeme und multimediale Lehrmaterialien die Lernbedingungen und fördern die aktive Teilnahme sowie die Entwicklung beruflicher Kompetenzen (Ruiz et al., 2006). Eine Studie von Sandberg (2022) zeigt darüber hinaus, dass Lehrkräfte digitale Technologien als hilfreich erachten, um den individuellen Anforderungen und Bedürfnissen der Lernenden auf verschiedenen Ebenen gerecht zu werden. Zudem können adaptive pädagogische Gesprächsagenten die Lehrkräfte bei der Erstellung personalisierter Materialien unterstützen, sodass sie sich auf andere Aspekte des Unterrichts konzentrieren können (Kasneji 2023). Auf diese Weise ist es möglich, den Unterricht individueller zu gestalten und auf unterschiedliche Förderbedürfnisse einzugehen. Darüber hinaus tragen präzises Feedback und die Anpassung der Lerninhalte an die individuellen Fähigkeiten der Lernenden zu einer effektiven Wissensbewertung bei (Minn 2022). Entscheidend ist, dass in einer personalisierten Lernumgebung die Lernziele und -inhalte, aber auch die Lernmethoden und das Lerntempo variieren können. Die Adaption der Inhalte kann dabei auf kognitiven Eigenschaften oder Verhaltensmerkmalen der Lernenden basieren, aber auch auf ihrem Lernstil, ihrem

Wissensstand oder ihren Lernpräferenzen (Nguyen und Nguyen 2023). Die Berücksichtigung individueller Unterschiede beim Lernen ist vor allem im Hinblick auf die Förderung akademischer Leistungen und die Entwicklung digitaler Kompetenzen von Lernenden von entscheidender Bedeutung (Bayly-Castaneda et al. 2024). Dabei müssen bei der Adaption von Lerninhalten alle für die Lernenden relevanten Merkmale berücksichtigt werden, denn nur dann kann die Personalisierung in allen Lernumgebungen erfolgreich sein (Tetzlaff et al. 2021). In diesem Zusammenhang nehmen Technologien wie Empfehlungssysteme, die auf maschinellem Lernen basieren, eine Schlüsselrolle ein, da sie die Adaption von Lerninhalten und somit eine effektivere Lernerfahrung ermöglichen (Desai 2020). Die Vielzahl an Einsatzmöglichkeiten von Technologien im Bildungsbereich verdeutlicht die Relevanz und Effektivität der Adaption von Lerninhalten. Sie ermöglichen eine präzise Anpassung der Lehrmethoden und Materialien, um sowohl den Lernprozess als auch die Kompetenzentwicklung zu optimieren und den Lernenden Fähigkeiten zu vermitteln, die in der modernen Arbeitswelt unerlässlich sind (Shihab et al. 2023). Im Bereich der Bildung wird KI bereits seit einigen Jahren zur Optimierung von Lernprozessen eingesetzt, z. B. um die Kursauswahl in Universitäten zu individualisieren oder Bildungslücken zu identifizieren und zu adressieren (Chen et al. 2020). So zeigt eine Studie von St-Hilaire et al. (2022), dass ein personalisiertes, aktives und KI-gestütztes Lernsystem die Lernergebnisse im Vergleich zu einem klassischen Lernsystem signifikant verbessert und dazu führt, dass Lernende eine höhere Abschlussquote und einen zwei- bis zweieinhalbfach höheren Lernzuwachs erzielen als Lernende, die ein klassisches Lernsystem ohne Adaption nutzen. Durch die Entwicklung der GKI, insbesondere der LLMs, haben sich neue innovative Wege zur Adaption von Lerninhalten eröffnet. Diese Methoden ermöglichen eine vollautomatische Anpassung, die deutlich über die Fähigkeiten früherer Ansätze hinausgeht. So zeigen Forschungsarbeiten, wie mithilfe von LLMs personalisiertes Feedback gegeben werden kann, das individuell auf das Wissen der Studierenden, z. B. zur mathematischen Problemlösekompetenz, abgestimmt ist (Jin et al. 2024), oder wie GKI als ergänzendes Bewertungssystem in Hochschulprüfungen eingesetzt werden kann (Karademir Coskun und Alper 2024). Zusätzlich haben Studien das breite Potenzial von LLMs für die Adaption von Lerninhalten beleuchtet. Ein Beispiel dafür ist die Arbeit von Leiker et al. (2023), die zeigt, dass LLMs hochwertige Lernmaterialien für Erwachsene generieren können, insbesondere wenn ein menschlicher Überprüfungsprozess zur Sicherstellung der inhaltlichen Qualität integriert wird. Auch in der Sprachlehre eröffnen LLMs neue Möglichkeiten. So wurde festgestellt, dass diese Modelle in der Lage sind, die Textschwierigkeit gezielt an die jeweilige Zielgruppe anzupassen, was insbesondere im Bereich der Sprachvermittlung eine bedeutende Rolle spielt (Caines et al. 2023).

Die Lesbarkeit von durch LLMs generierten Texten wurde in einer Untersuchung von Rooein et al. (2023) analysiert, wobei deutliche Unterschiede in der Eignung für verschiedene Bildungsniveaus festgestellt wurden. Dies verdeutlicht den Bedarf, LLMs gezielt auf bestimmte Zielgruppen hin zu optimieren.

Experimente von Gobara et al. (2024) zeigen, dass LLMs die Schwierigkeit von Texten automatisch anpassen können und in bestimmten Bereichen, wie dem Abstimmen der Textschwierigkeit auf Eingaben, sogar menschliche Leistungen über-

treffen. Diese Fortschritte ermöglichen es, Lerninhalte effektiver und individueller zu gestalten, während gleichzeitig die Effizienz der Lehrmaterialeerstellung erhöht wird. Bayly-Castaneda et al. (2024) weisen darauf hin, dass Forschung zur Personalisierung von Lernen durch hybride Lösungen, also Ansätze, die Mensch und Maschine kombinieren, bisher nur in wenigen Ländern betrieben wird. Besonders relevant sind dabei Anwendungen, die auf adaptivem Lernen und LLMs basieren, wie zum Beispiel mobile Lernumgebungen, Lernspiele, soziale Netzwerke, Augmented Reality und Massive Open Online Courses. Bewertungs- und Feedbackprozesse zu diesen Technologien (Ifelebuegu 2023) bleiben ebenfalls ein wichtiger Aspekt. Die geografische Verteilung der Forschung, basierend auf der Herkunft des Erstautors, ist stark auf China konzentriert. Es folgen die USA, Indien und Deutschland mit einem Verhältnis von fünf zu eins zu China. Afrika und Lateinamerika spielen bislang eine untergeordnete Rolle. Dies deutet darauf hin, dass das Potenzial für personalisierte KI-gestützte Lernlösungen global noch nicht ausgeschöpft ist, obwohl kulturelle Unterschiede und Herkunftskontexte das Lernverhalten beeinflussen können. In Deutschland besteht somit im Vergleich zu China noch deutlicher Aufholbedarf. Um den Anschluss nicht zu verlieren und die Chancen solcher Technologien voll zu nutzen, ist mehr Forschung zu KI-gestützten, personalisierten Lernumgebungen dringend erforderlich (Bayly-Castaneda et al. 2024). Folglich befinden sich viele der Forschungsarbeiten gegenwärtig noch in den experimentellen und theoretischen Entwurfsphasen. In Zukunft ist jedoch eine Zunahme der Evaluierungen der Effektivität der Anwendungen zu erwarten, da diese Forschungsrichtung zunehmend an Reife gewinnt.

Hier setzt dieser Artikel an: Wir untersuchen systematisch, wie LLMs mithilfe eines Prompt-Engineering-Ansatzes bestehende Inhalte auf individuelle Interessen zuschneiden können, um Lerntexte für die Einzelnen motivierender zu gestalten. Damit greifen wir eine Lücke auf, die sich in der bisherigen Forschung abzeichnet: Trotz vielversprechender GKI- und Adaptionansätze liegt der Schwerpunkt bisher vor allem auf der Anpassung von Textschwierigkeiten (Caines et al. 2023; von Gobarra et al. 2024) oder personalisiertem Feedback (Jin et al. 2024). Die spezifische Adaption basierend auf individuellen Interessen mit LLMs wurde hingegen, insbesondere im deutschen Raum (Bayly-Castaneda et al. 2024), bislang kaum empirisch untersucht. Die Funktionsweise von LLMs basiert auf dem Prinzip der statistischen Sprachmodellierung. Dazu wird eine Transformer-Architektur verwendet, welche die Beziehungen zwischen den Wörtern einer Sequenz erfasst. Die Eingabesequenz wird auf Grundlage ihrer Folge in eine andere Sequenz (die Ausgabe) übersetzt (Vaswani 2017). Bislang wurden LLMs beispielsweise von Huang und Rust (2024) in der Kundenbetreuung eingesetzt, um die Qualität emotional belasteter Kundeninteraktionen zu verbessern. Ihr Ansatz beschreibt eine Entwicklung von der genauen Emotionserkennung über empathische Reaktionen und emotionalem Management bis hin zum Aufbau einer emotionalen Verbindung und deutet auf eine wesentliche Verbesserung gegenüber früheren Technologien wie Chatbots hin, die über eingeschränktere emotionale Fähigkeiten verfügen. Zusätzlich zum Eingehen auf die persönlichen Bedürfnisse der Nutzenden kann die inhaltliche Qualität der Antworten in Zukunft noch weiter verbessert werden, beispielsweise kann die Retrieval Augmented Generation-Technologie die Antworten des LLMs durch die Integration

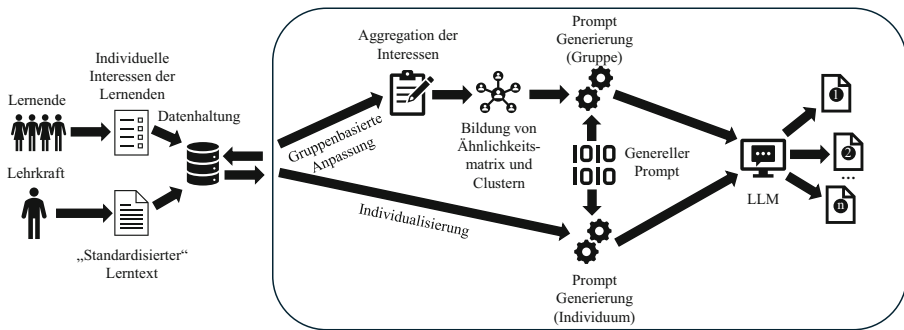


Abb. 1 Ablauf der Adaption

von Informationen aus kuratierten Quellen präzisieren. Dieser Prozess verbessert die Fähigkeit des Modells, spezialisierte und aktuelle Informationen zu liefern (Lewis et al. 2020). An der Schnittstelle zur Bildung wurden GKI wie LLMs eingesetzt, um zu untersuchen, inwieweit Lehrkräfte im Schulalltag unterstützt werden können, etwa durch die Erstellung neuer Aufgaben (Hein et al. 2024). Vor diesem Hintergrund baut diese Studie auf den Erkenntnissen dieser Arbeiten auf, indem sie den Fokus auf die automatische Anpassung bereits bestehender Aufgaben legt und untersucht, inwieweit sich diese an die individuellen Interessen der einzelnen Lernenden angleichen lassen.

3 Interessensbasierte Adaption von Lerntexten durch Künstliche Intelligenz

3.1 Interessenerhebung und Prompt-Generierung

Abb. 1 zeigt den Prozess der Adaption von Lerninhalten durch den Einsatz von GKI. Um eine Adaption der Lerntexte zu erreichen, werden zunächst die Interessen der Lernenden durch eine Vorab-Befragung ermittelt. In diesem Zusammenhang wurden den Befragten die folgenden Fragen gestellt:

1. Welche Hobbys und Freizeitaktivitäten üben Sie regelmäßig aus (z.B. Fantasy lesen, Malen, Kochen, Fußball)?
2. Welche Aktivitäten verfolgen/gucken Sie aktiv (z.B. Fußball)?
3. Welche wissenschaftlichen oder akademischen Bereiche interessieren Sie besonders?
4. Welche spezifischen Themen oder Fragestellungen ziehen Ihr Interesse besonders an? Diese können kultureller, technologischer oder sozialer Natur sein.

Hierbei konnten die Interessen frei angegeben werden, ohne Unterscheidung zwischen beispielsweise kognitiven oder körperlichen Interessen. Die erhobenen Daten werden daraufhin automatisiert verarbeitet und in einer Datenbank gespeichert. Hierauf aufbauend wurden ähnliche Interessen zusammengefasst, z.B. „Fitnessstudio“,

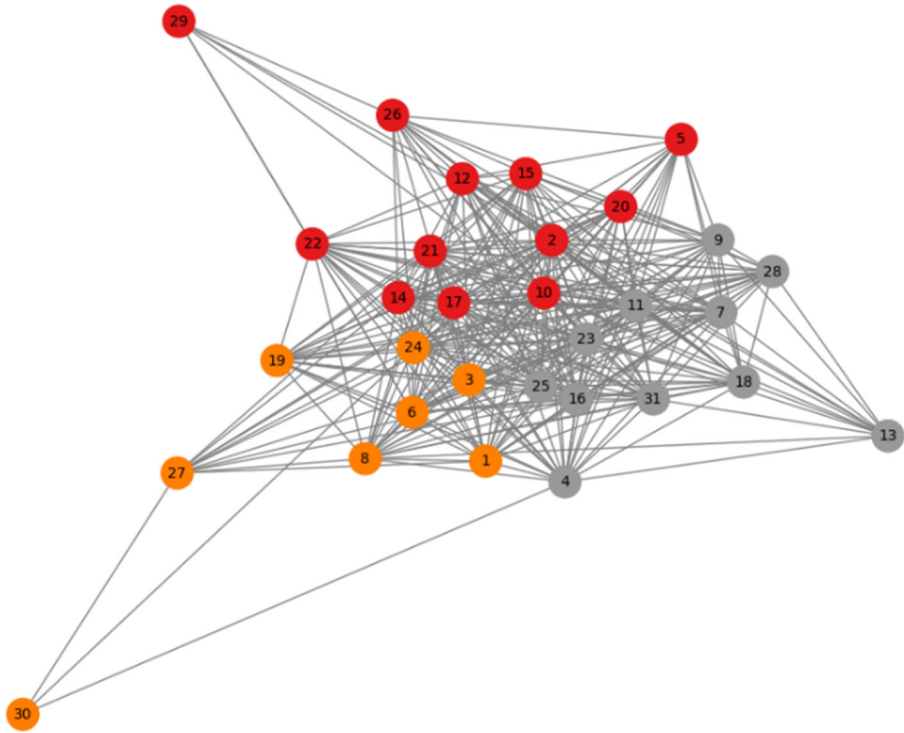


Abb. 2 Netzwerkgraph

„Gym“ und „körperliche Fitness“ zu „Fitness“, um die Gruppenanpassung der Lern-
 texte zu ermöglichen. Basierend auf diesen Interessensgruppen wurde eine Ähnlich-
 keitsmatrix erstellt, wobei Teilnehmende mit gleichen Interessensgruppen jeweils
 einen Ähnlichkeitspunkt erhielten. Mithilfe dieser Ähnlichkeitsmatrix wurden Cluster
 mittels des K-Means-Algorithmus gebildet. K-Means wurde für das Clustering
 gewählt, da er ein etabliertes Verfahren ist, das Datenpunkte basierend auf Ähnlich-
 keiten gruppiert. Der Algorithmus teilt die Daten in eine vorab definierte Anzahl von
 Clustern auf, sodass die Punkte innerhalb eines Clusters möglichst ähnlich und zwi-
 schen den Clustern möglichst unterschiedlich sind (MacQueen 1967). Diese Cluster
 werden abschließend mit einem Netzwerkgraphen visualisiert (Abb. 2).

Anschließend wurde ein System-Prompt definiert, der für das genutzte LLM
 (GPT-4) festlegt, welche Art von Input verarbeitet wird, wie dieser verarbeitet wird
 und was der erwartete Output ist. Der Prompt wurde in einem iterativen Prozess
 entwickelt, bei dem er zunächst mit mehreren Nutzern und deren Interessen getestet
 wurde. Die Autoren prüften anschließend die generierten Ergebnisse auf Validität
 und nahmen gezielte Anpassungen vor. In mehreren Iterationen wurden Formulier-
 ungen gestrichen oder ergänzt, um die Ausgaben des LLM zu optimieren. Beispiels-
 weise wurde die Formulierung „*Stelle sicher, dass dein neuer Text [...] sich eng am
 Originalinhalt orientiert*“ sowie „*Der angepasste Text sollte in Länge und Struktur
 so nah wie möglich am Original bleiben*“ ergänzt, da das Modell initial dazu neigte,

neue Konzepte zu erfinden und die ursprünglichen Inhalte durch diese unnötig zu erweitern. Um zu verhindern, dass das Modell zusätzliche Metakommentare wie Rechtfertigungen oder Zusammenfassungen der generierten Inhalte ausgibt, wurde später die Anweisung „Gib nur den angepassten Text aus“ hinzugefügt.

Der final entwickelte Prompt lautete: „Du bist ein Modell zur Anpassung von Lerntexten an die spezifischen Interessen der Lernenden. Deine Aufgabe ist es, den Lerntext, insbesondere die Beispiele, so anzupassen, dass er den Interessen der Lernenden entspricht. Verwende dazu eines der angegebenen Interessen. Stelle sicher, dass dein neuer Text das Verständnis der grundlegenden Konzepte fördert, kohärent ist und sich eng am Originalinhalt orientiert. Der angepasste Text sollte in Länge und Struktur so nah wie möglich am Original bleiben. Gib nur den angepassten Text aus.“

Zusätzlich zum System-Prompt wird ein genereller Prompt definiert, der die Struktur für die Weitergabe des Originaltextes und einer Liste der Interessen an das System festlegt: „Lerntext: [Plaintext], Interessen: [Interessen]“

Nach Festlegung des Prompts wird der zu modifizierende Lerntext eingelesen, in Plaintext konvertiert und in der Datenbank abgelegt. Das System erstellt daraufhin aus den drei Komponenten – genereller Prompt, Lerntext und Interessen des Lernenden – einen spezifischen Prompt für jeden einzelnen Lernenden bzw. für jede Lerngruppe, indem die Platzhalter in dem generellen Prompt durch den entsprechenden Lerntext und eine Liste der Interessen des Lernenden bzw. die Top 5 Interessen der Gruppe ersetzt werden. Die Interessen wurden dabei ohne direktes Matching zu den Themen an das LLM weitergegeben, um eine vollautomatische Anpassung ohne menschlichen Einfluss zu ermöglichen. Dieses Vorgehen überließ die Verantwortung, ein passendes Interesse für das Themengebiet zu wählen, dem LLM.

Die spezifischen, individualisierten Prompts werden daraufhin für jeden Lernenden und für jeden Lerntext ausgeführt. Das generierte Ergebnis wird ebenfalls in der Datenbank gespeichert. Abb. 3 zeigt einen angepassten Inhalt gemäß den Interessen Fantasy-Literatur, Fotografie, Magic the Gathering, Warhammer etc., wobei das LLM sich auf das Interesse Magic the Gathering konzentriert.

Standard-Text

Ein System kann als eine Menge von Elementen definiert werden, die durch ihre Eigenschaften charakterisiert sind und in spezifischen Beziehungen zueinanderstehen. Diese Definition lässt sich am Beispiel eines Unternehmens verdeutlichen. Die Elemente eines Unternehmens umfassen Ressourcen wie Mitarbeiter, Technologie und Kapital. Jedes dieser Elemente besitzt spezifische Eigenschaften, die ihre Funktion und Rolle im Unternehmenskontext definieren, wie etwa die Qualifikationen der Mitarbeiter oder die Kapazität der Technologie. Die Beziehungen zwischen diesen Elementen manifestieren sich in der Organisationsstruktur und den Arbeitsabläufen, die koordinieren, wie diese Ressourcen zur Erreichung der Unternehmensziele zusammenwirken. Ein System ist somit nicht nur eine Sammlung von Elementen, sondern eine dynamische Einheit, die durch die Interaktionen zwischen ihren Bestandteilen gekennzeichnet ist. Diese systemische Sicht ermöglicht es, komplexe Strukturen und Prozesse in Unternehmen effektiv zu analysieren und zu gestalten.

Individualisierter-Text

Ein System kann als eine Menge von Elementen definiert werden, die durch ihre Eigenschaften charakterisiert sind und in spezifischen Beziehungen zueinanderstehen. Lassen Sie uns dies am Beispiel eines Magic the Gathering (MtG) Kartenspiels verdeutlichen. Die Elemente in einem MtG-Spiel umfassen Ressourcen wie Karten, Mana und Lebenspunkte. Jedes dieser Elemente besitzt spezifische Eigenschaften, die ihre Funktion und Rolle im Spielkontext definieren, wie etwa die Fähigkeiten der Karten oder die Menge an verfügbaren Mana. Die Beziehungen zwischen diesen Elementen manifestieren sich in den Spielregeln und Strategien, die koordinieren, wie diese Ressourcen zur Erreichung des Spielziels - den Gegner zu besiegen - zusammenwirken. Ein MtG-Spiel ist somit nicht nur eine Sammlung von Karten, sondern eine dynamische Einheit, die durch die Interaktionen zwischen ihren Bestandteilen gekennzeichnet ist. Diese systemische Sicht ermöglicht es, komplexe Strukturen und Prozesse in einem MtG-Spiel effektiv zu analysieren und zu gestalten.



Abb. 3 Beispiel für einen individualisierten Lerntext

3.2 Evaluation der erstellten Lerntexte

Um den Einfluss der GKI auf die Adaption von Lerntexten zu untersuchen, wurde ein Experiment durchgeführt. Ziel dieses Experiments ist es zu ermitteln, wie sich die vier Textarten in Bezug auf Engagement, Verständlichkeit, wahrgenommene Nützlichkeit und Gesamteindruck unterscheiden. In die Evaluation wurden insgesamt 12 verschiedene Texte einbezogen, die Themen wie Systeme, die Principal-Agent-Theorie oder die Theorie der sozialen Identität erläutern. Die Teilnehmenden erhielten jeweils vier STL sowie vier ILT. Bezüglich der Anpassung auf die Interessen der Gruppe wurden die Teilnehmenden ein weiteres Mal aufgeteilt. Der eine Teil der Teilnehmenden erhielt vier GLT, der andere Teil vier FLT. Um eine breite Datenbasis zu gewährleisten, wurde unter den Teilnehmenden rotiert, wer welchen Text wie angezeigt bekam. Jedes Thema wurde so jeweils einem Drittel der Teilnehmenden als SLT, einem weiteren Drittel als ILT und dem restlichen Drittel als GLT oder FLT präsentiert. Um Reihenfolgeeffekte zu minimieren, erfolgte die Darstellung der Texte in zufälliger Reihenfolge. Auf der Grundlage der Texte wurde ein Fragebogen mit etablierten Skalen auf einer Likert-Skala von 1 (stimme überhaupt nicht zu) bis 7 (stimme voll und ganz zu) verwendet, um verschiedene Aspekte der Textwirkung zu messen (Tab. 1). Die Reliabilität der drei gemessenen Konstrukte (Engagement, Verständlichkeit und wahrgenommene Nützlichkeit) ist mit Cronbachs Alpha Werten zwischen 0,88 und 0,91 gut.

Der Gesamteindruck des Textes wurde ebenfalls mit der Likert-Skala abgefragt. Zusätzlich wurden nach jedem Text zwei qualitative Fragen gestellt, was den Teilnehmenden am spezifischen Lerntext besonders gefallen bzw. missfallen hat und was sie ändern würden. Am Ende der Evaluation wurden drei abschließende qualitative Fragen gestellt: Welche Gemeinsamkeiten hatten die Texte, die den Teilnehmern am besten gefallen haben, welche Gemeinsamkeiten die Texte hatten, die den Teilnehmern am wenigsten gefallen haben und ob sie weitere Gedanken teilen möchten. Im Anschluss an die Evaluation wurden die gesammelten Daten mittels einer Varianzanalyse analysiert, um signifikante Unterschiede zwischen den Bewertungen der SLT, ILT, GLT und FLT zu identifizieren. Mithilfe des Tukey HSD (Tukey 1949) wurden anschließend spezifische Unterschiede zwischen den Gruppen überprüft.

Tab. 1 Konstrukte der Evaluation

Konstrukt	Beispiel-Item	Etablierte Skala
Engagement	Wenn ich den Text lese, fühle ich mich motiviert	Schaufeli et al. 2002
Verständlichkeit	Der Text ist starr und unflexibel gestaltet, was das Verständnis erschwert	Davis 1989
Wahrgenommene Nützlichkeit	Mit diesem Text kann ich Lerninhalte schneller begreifen	Davis 1989

3.3 Stichprobenzusammensetzung

Im Rahmen der Erhebung der Interessen wurden insgesamt 31 Personen befragt, welche sich auf Basis eines Aufrufs des Forschungsteams zu einem Experiment zur Adaption in der Bildung angefragt worden waren, um an der Studie teilzunehmen. Dabei wurde keine Selektion vorgenommen, jede interessierte Person, die sich freiwillig meldete, konnte an der Studie teilnehmen. Der Aufruf zur Teilnahme erfolgte unter Berücksichtigung verschiedener Tätigkeitsbereiche. Hierbei wurde neben dem Wissensstand und dem Anwendungsniveau mit GKI ebenso deren jeweilige Tätigkeit berücksichtigt. Die Stichprobe setzt sich demnach aus Personen zusammen, die im Bildungsbereich aktiv sind, in der (Wirtschafts-)Informatik arbeiten sowie aus Personen, die dem sozialwissenschaftlichen Bereich zuzuordnen sind. Dabei wurden für die Gruppeninteressen Cluster nach Abb. 2 gebildet. Im Anschluss an die Interessenerhebung haben 26 von 31 Personen den zweiten Teil der Studie beantwortet und die Texte bewertet. Von diesen Personen hat eine den Aufmerksamkeitstest nicht bestanden, sodass sich die Stichprobe, die in die Auswertung mit einbezogen werden konnte, aus 25 Personen zusammensetzt. Die Teilnehmenden der Studie waren zwischen 23 und 52 Jahren, von denen 10 weiblich und 15 männlich waren. Die mediane Zeit, um die Evaluation abzuschließen, betrug 61 min. Insgesamt wurden 4800 quantitative und 358 qualitative Datenpunkte erhoben.

4 Ergebnisse im Vergleich: Einfluss individualisierter Lerntexte auf das Lernerlebnis

Die Ergebnisse (Tab. 2) zeigen, dass die ILT höhere Werte bei den drei Konstrukten Engagement, Verständlichkeit und Gesamteindruck erreichten, als die SLT. GLT erhielten niedrigere Bewertungen im Engagement, Verständlichkeit und Wahrgenommene Nützlichkeit als SLT, aber eine bessere Bewertung im Gesamteindruck. FLT schnitten in Engagement, Verständlichkeit und Wahrgenommene Nützlichkeit schlechter ab als SLT.

Die Varianzanalyse (Tab. 3) ergab signifikante Unterschiede auf den Skalen Engagement ($p < 0.05$) sowie eine signifikante Tendenz bei der Skala Wahrgenommene Nützlichkeit ($p = 0.07$). Beim Gesamteindruck ($p = 0.74$) und bei der Verständlichkeit ($p = 0.08$) konnten keine signifikanten Unterschiede festgestellt werden. Basierend auf der Tukey HSD zeigen sich bei der Skala Engagement signifikante Unterschiede zwischen den ILT und den GLT ($p < 0.05$). Auf der Skala Wahrgenommene Nütz-

Tab. 2 Mittelwerte der Ergebnisse der Evaluation

	SLT	ILT	GLT	FLT
Engagement	4,71	4,87	4,24	4,46
Verständlichkeit	4,54	4,73	4,15	4,15
Wahrgenommene Nützlichkeit	5,57	5,35	5,26	4,90
Gesamteindruck	3,82	4,01	4,11	3,89

Tab. 3 Ergebnisse ANOVA Engagement

Engagement	SLT			ILT			GLT			FLT		
	ILT	GLT	FLT	SLT	GLT	FLT	SLT	ILT	FLT	SLT	ILT	GLT
Sig	0,84	0,20	0,77	0,84	0,04*	0,37	0,20	0,04*	0,87	0,77	0,37	0,07
SD	0,20	0,24	0,25	0,20	0,24	0,25	0,24	0,24	0,28	0,25	0,25	0,28

Tab. 4 Ergebnisse ANOVA Wahrgenommene Nützlichkeit

Wahrgenom- mene Nützlich- keit	SLT			ILT			GLT			FLT		
	ILT	GLT	FLT	SLT	GLT	FLT	SLT	ILT	FLT	SLT	ILT	GLT
Sig	0,07	0,57	0,04*	0,70	0,98	0,28	0,57	0,98	0,57	0,04*	0,28	0,57
SD	0,20	0,24	0,26	0,20	0,24	0,26	0,24	0,24	0,28	0,26	0,26	0,28

lichkeit ist eine Tendenz zu signifikanten Unterschieden zwischen den Texten der SLT und den FLT ($p < 0.07$) vorhanden (Tab. 4).

Neben der Auswertung der statistischen Skalen wurden die offenen Fragen vom Forschungsteam qualitativ analysiert und anschließend mithilfe von Codes quantifiziert. Die Analyse des Gesamteindrucks über alle gezeigten Texte hinweg zeigte, dass die Texte, die den teilnehmenden Personen am besten gefallen haben, alle die Gemeinsamkeit hatten, dass sie den Interessen der Lernenden entsprachen ($n = 19$), die Theorie anhand von Beispielen verständlich dargestellt ($n = 11$) und eine verständliche und nachvollziehbare Sprache verwendet wurde ($n = 8$). Im Gegensatz dazu wurden bei den Texten, die am wenigsten gefallen haben, Gemeinsamkeiten gefunden, die die Schlüssigkeit der Texte sowie das Vorhandensein von Logikfehlern ($n = 12$) beinhalteten. Zudem wurde angemerkt, dass bei den Texten der Bezug zu den Interessen ungünstig gewählt war oder die Interessen nicht den zuvor angegebenen Interessen entsprachen ($n = 7$). Des Weiteren wurde beobachtet, dass die Texte teilweise sehr verschachtelt ($n = 6$) und mit Fachwörtern und Fachthemen versehen seien, was das Verständnis erschwere ($n = 4$). Zudem wurde in einigen Fällen eine zu statische Textgestaltung bemängelt, die den Eindruck erweckt, als seien die Texte maschinell erstellt worden ($n = 3$). Als Verbesserungsvorschlag wurde erwähnt, dass der Textaufbau verbessert werden kann, der Wissensstand der Lernenden berücksichtigt werden sollte und die Möglichkeit bestehen sollte, Rückfragen zu stellen. Dies könne insbesondere im Hinblick auf die Verwendung von Fachwörtern von Vorteil sein. Eine Analyse der einzelnen Texte zeigt, dass die Bewertungen der Teilnehmenden nicht konsistent sind. Es kann festgestellt werden, dass nicht alle Interessen geeignet sind, einen Lerntext verständlich zu erklären. So eignet sich beispielsweise das Interesse Kunstgeschichte nicht, um Systeme zu erläutern. Jedoch konnte der Systembegriff verständlich vermittelt werden bei den Interessen Chor, Kochen bzw. Backen in Form von einem Rezept oder anhand des Kartenspiels Magic the Gathering. Bei einfachen Texten, die beispielsweise die Begriffe Effizienz und Effektivität erklären, hielten die Teilnehmenden die Einbeziehung von Interessen nicht für notwendig.

5 Potenziale und Herausforderungen der KI-getriebenen Adaption von Lernertexten

In der Studie wurde das Potenzial von GKI zur Adaption von Lernertexten untersucht, wobei verschiedene Gruppen (SLT, ILT, GLT, FLT) verglichen wurden. Die Ergebnisse geben erste Hinweise, dass Adaption mit GKI grundsätzlich geeignet sein kann, Lernertexte zu verbessern, vorausgesetzt, die individuellen Interessen der Lernenden werden berücksichtigt und in den Kontext der Inhalte gesetzt. Obwohl ILT in den etablierten Skalen zu Engagement, Gesamteindruck und Verständlichkeit besser abschneiden als SLT, ist der Unterschied nicht statistisch signifikant. Schlechtere Bewertungen in den ILT treten auf, wenn die Beispiele nicht zu den Interessen der Lernenden passen. Dieses Phänomen wurde in den Gruppen GLT und FLT noch verstärkt, da die ungeeigneten Beispiele in diesen seltener oder gar nicht durch für die Lernenden interessante Beispiele kompensiert werden konnten. Wenn die Interessen des Lernenden nicht getroffen werden, verliert der Lernertext zusätzlich zu dem bewusst gewählten Beispiel des SLT an Identifikation und wird schlechter bewertet. Die Ergebnisse der Studie legen nahe, dass, wenn eine Adaption nach Gruppen gewünscht ist, eine mögliche Lösung darin bestehen könnte, die Gruppen enger zu fassen, sodass jede Gruppe nur Lernende mit bestimmten Interessen umfasst.

Unsere Studie zeigt erste Ansätze auf, wie Bildung mithilfe von GKI verbessert werden kann. Dabei bekräftigen unsere Ergebnisse frühere Studien von St-Hilaire et al. (2022), die auf die Bedeutung eines personalisierten, aktiven und KI-gestützten Lernsystems hinweisen. Differenzierung durch Adaption wurde in der Vergangenheit immer wieder als Schlüssel zur Erreichung von qualitativ hochwertiger Bildung genannt (Altrichter et al. 2009; McDonald Connor et al. 2009; Raggl 2023). Technologien wurden dabei als Schlüsselaspekt für die Umsetzbarkeit erachtet, um Lehrmethoden noch dynamischer zu gestalten (Karpenko 2019). In der Vergangenheit wurde das Erreichen eines solchen Individualisierungsgrades jedoch als nicht realisierbar angesehen, da den Lehrkräften die Zeit fehlte, auf die Interessen und Wünsche jedes einzelnen Lernenden einzugehen (Tomlinson 2001). Technologien eröffnen nunmehr die Möglichkeit einer Adaption, indem sie Lehrkräften eine Berücksichtigung individueller Anforderungen und Bedürfnisse der Lernenden auf verschiedenen Ebenen erlauben (Sandberg 2022). Dies ist darauf zurückzuführen, dass Technologien die Lehrkräfte bei der Erstellung personalisierter Materialien unterstützen, sodass sich diese auf andere Aspekte des Unterrichts konzentrieren können (Kasneci 2023).

Die Ergebnisse unserer Studie zeigen, dass die GKI zwar das Potenzial bietet, diese Adaption zu realisieren, jedoch nicht mit dem in dieser Studie verwendeten unüberwachten Ansatz. So ist eine Überprüfung der Texte vor der Bearbeitung durch die Lernenden in diesem Zusammenhang unerlässlich. Wie bereits in früheren Studien festgestellt, mangelt es der GKI an didaktischem Verständnis (Hein et al. 2024). Sie ist zwar in der Lage, die Texte an die Interessen anzupassen, scheint dabei aber willkürlich eines der angegebenen Interessen auszuwählen, ohne zu berücksichtigen, ob dieses zum Inhalt des Textes passt. Das Ergebnis ist ein angepasster Text, der für den Lernenden entweder sehr ansprechend ist oder zusammenhangslos erscheint. Hier zeigt sich auch die Tendenz der GKI – obwohl im Prompt anders angegeben – in Texten mit mehreren Konzepten unterschiedliche Interessen zusammenzuführen.

Dies verschlechtert die Qualität der angepassten Lerntexte zusätzlich. GKI ist bisher noch nicht auf dem Entwicklungsstand, um unbeaufsichtigt, ohne die Kontrolle durch eine Lehrkraft, didaktische Aufgaben zu übernehmen. Ein engeres Zusammenspiel von Technologie und Didaktik könnte dieses Problem adressieren, um bessere Ergebnisse zu liefern. So könnte eine Lehrkraft vorab auswählen, zu welchen Interessen ein von ihr eingegebener Standardlernertext angepasst werden soll, sofern sie einen Zusammenhang zwischen Thema und Interesse sieht. Im Anschluss daran kann die Generierung des Textes erfolgen und von der Lehrkraft beurteilt werden, ob sie diesen so übernimmt oder eine neue Textgenerierung anstößt. Wie bereits von Tomlinson (2001) für Ansätze, die in der Vergangenheit genutzt wurden, beschrieben, ist auch dieser Adaptionsansatz mit einem Mehraufwand für die Lehrkraft verbunden, jedoch wird dieser Aufwand deutlich geringer sein, als einen Text selbstständig anzupassen.

Die Ergebnisse zeigen somit erste Potenziale, aber auch Herausforderungen bei der Nutzung von GKI zur Verbesserung von Lernmaterialien. Die Möglichkeit, Lerninhalte so zu optimieren, dass sie auf individuelle Nutzerprofile zugeschnitten sind, könnte den Lehrenden einen Weg eröffnen, das Engagement und die Bindung der Lernenden signifikant zu erhöhen, sofern dies durch eine durchdachte Kombination von Technologie und didaktischer Expertise unterstützt wird. Es sei jedoch darauf verwiesen, dass mit der Nutzung von GKI potenzielle Risiken einhergehen, die es zu berücksichtigen gilt. Besonders kritisch sind Aspekte des Datenschutzes und der Verarbeitung personenbezogener Daten, wie etwa die Angabe von Namen, Informationen zum Förderstatus oder potenziellen Leistungsdefiziten (Hein et al. 2024). Darüber hinaus birgt die zugrunde liegende Trainingsdatenbasis generativer KI das Risiko algorithmischer Verzerrungen, die zu unfairen oder diskriminierenden Ergebnissen führen können – beispielsweise aufgrund von Geschlecht, ethnischer Herkunft oder Behinderungsstatus. Dies unterstreicht die Notwendigkeit einer sorgfältigen Überprüfung der generierten Ergebnisse, um derartigen Verzerrungen entgegenzuwirken (Kleinberg 2018; Noble 2018).

Eine weitere Herausforderung besteht darin, dass die derzeitige Konzentration der Forschungs- und Entwicklungsaktivitäten in wenigen Ländern das Risiko algorithmischer Verzerrungen erhöht und gleichzeitig die digitale Kluft verstärken kann. Diese Entwicklung macht deutlich, wie wichtig ein sorgfältiger und ethisch fundierter Ansatz bei der Implementierung dieser Technologien ist (Bayly-Castaneda et al. 2024).

Aus pädagogischer Sicht wird zudem kontrovers diskutiert, ob adaptive Systeme – darunter auch GKI – die Autonomie und Entscheidungsfreiheit von Lernenden einschränken können. Um dieser Problematik zu begegnen, wird die Integration eines hybriden Mensch-KI-Ansatzes im Bildungsbereich als unverzichtbar angesehen. Dabei sollte den Praktikerinnen und Praktikern die Möglichkeit gegeben werden, die Lehrinhalte flexibel anzupassen, etwa in Form von Co-Teaching-Szenarien, um die Vorteile der KI zu nutzen und zugleich die Kontrolle durch menschliche Akteure sicherzustellen (Holstein und Alevan 2022; Schmid et al. 2021).

Die Anpassung von Lernmaterialien kann über die Steigerung der Nutzerzufriedenheit hinaus auch kommerzielle Vorteile bieten, beispielsweise durch eine erhöhte Nutzerakzeptanz und eines stärkeren Engagements. Gleichzeitig trägt die Adaption

von Lerninhalten mittels GKI zur Demokratisierung der Bildung bei, indem sie deren Zugänglichkeit erhöht und hilft, Bildungsungleichheiten zu überwinden. Dies ist entscheidend, um diverse Gruppen aus Lernenden auf die Herausforderungen moderner Wissensökonomien vorzubereiten und das Abbruchrisiko in Schlüsseldisziplinen wie der IT zu minimieren. GKI-unterstützende Lernprozesse haben das Potenzial, sowohl akademisch bereichernd als auch persönlich relevant zu sein. Hierbei spielt die Wirtschaftsinformatik eine Schlüsselrolle, da sie als Schnittstelle zwischen Didaktik und Technologie wirkt und innovative Bildungsansätze ermöglicht.

6 Fazit und Ausblick auf zukünftige Forschung

Unsere Studie zeigt das Potenzial der GKI für die Adaption von Lerntexten. Dazu bewerteten 25 Personen jeweils 12 Texte, die entweder für sie selbst, für ihre Gruppe, für eine andere Gruppe oder gar nicht adaptiert worden waren. Die Ergebnisse geben erste Hinweise darauf, dass individualisierte Texte in den Bereichen Engagement, Verständlichkeit und Gesamteindruck tendenziell besser bewertet wurden als die Texte der anderen Gruppen, jedoch sind diese Ergebnisse nur teilweise statistisch signifikant. Probleme mit den ILT traten auf, da einige von ihnen auf der Grundlage von Interessen ungeeignete Beispiele verwendeten, die von den Lernenden abgelehnt wurden. Eine Adaption auf Gruppenebene ist nicht zu empfehlen, da die GLT trotz kleiner Gruppengröße schlechter bewertet wurden als die SLT. Die GLT haben dieselben Nachteile wie die ILT, können aber zusätzlich das Interesse der Lernenden verfehlen. Die FLT wurden ebenfalls schlechter bewertet als die SLT, was die Gefahren einer falsch durchgeführten Adaption verdeutlicht. Eine unpassende Anpassung kann somit die Zufriedenheit mit dem Lernmaterial senken. Um Probleme mit den Lerntexten zu adressieren, wurde seitens der Teilnehmenden die Empfehlung ausgesprochen, zusätzlich zu den individuellen Interessen auch den bisherigen Wissensstand in den Adaptionsprozess mit einzubeziehen. Dieses Vorgehen gilt es, in zukünftiger Forschung umzusetzen und zu evaluieren. Weiterhin sollte es ermöglicht werden, Rückfragen zum Text zu stellen. Dieser Ansatz erfordert den Einsatz von GKI zur Lernbegleitung bspw. durch Chatbots, auch hier muss ein Prozess entwickelt und evaluiert werden. Die vorliegende Studie ist mit gewissen Einschränkungen verbunden. Die Teilnehmenden waren zwischen 23 und 52 Jahren alt, sodass lediglich der Einfluss auf Erwachsene gemessen werden konnte. Zudem betrug die Stichprobe lediglich 25 Personen, was die Generalisierbarkeit der Ergebnisse begrenzt. Zukünftige Studien sollten auf einer breiteren und vielfältigeren Datenbasis aufbauen, die eine größere Anzahl an Interessen und demografischen Merkmalen umfasst. Weiterhin ist eine detaillierte Analyse der Wechselwirkungen zwischen den Lerninhalten und den individuellen Interessen der Lernenden erforderlich, um zu ermitteln, wie generative KI am effektivsten für die Adaption von Bildungsinhalten eingesetzt werden kann. Ein weiterer kritischer Aspekt dieser Studie liegt in der Nutzung eines einzigen Modells (GPT-4) zur Anpassung der Lerntexte. Obwohl GPT-4 zum Zeitpunkt der Studie eines der leistungsstärksten und präzisen Modelle war und daher eine hohe Aussagekraft vermutet werden kann, könnten andere Modelle zu abweichenden Ergebnissen führen. Darüber hinaus ist es möglich,

dass zukünftige Entwicklungen neuer LLMs oder speziell auf die Anpassung von Lerntexten trainierte Modelle (z. B. durch Finetuning) die Präzision und Effektivität weiter verbessern. Zukünftige Forschung sollte daher den Einfluss verschiedener LLMs sowie deren Anpassung untersuchen, um die Potenziale optimierter und spezialisierter Modelle im Bildungsbereich besser zu verstehen. Schließlich ist unser Ansatz maßgeblich von dem definierten Prompt abhängig. Auch wenn dieser iterativ entwickelt und verbessert wurde, könnte ein anderer Prompt zu besseren Ergebnissen führen. Insgesamt sehen wir trotz dessen großes Potenzial in der Nutzung von GKI zur Verbesserung von Lernprozessen, empfehlen jedoch einen Human-in-the-Loop-Ansatz z. B. durch die Überwachung der GKI durch eine Lehrkraft, um eine effiziente Verbindung zwischen Technologie und Didaktik zu gewährleisten. Zukünftige Forschungsarbeiten sollten daher untersuchen, wie eine von einer Lehrkraft kontrollierte Verwendung von GKI zur Adaption gestaltet werden kann und inwiefern sie zu verbesserten Ergebnissen führt.

Funding Open Access funding enabled and organized by Projekt DEAL.

Einhaltung ethischer Richtlinien

Interessenkonflikt D. Stattkus, L. Hein und O. Thomas geben an, dass kein Interessenkonflikt besteht.

Ethische Standards Für diesen Beitrag wurden von den Autor/-innen keine Studien an Tieren durchgeführt. Die Evaluation mit Menschen erfolgte unter Berücksichtigung der geltenden ethischen Richtlinien. Die Teilnehmenden nahmen freiwillig an der Bewertung der Lerntexte teil und gaben ihre informierte Einwilligung. Belastende oder invasive Maßnahmen wurden nicht durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Altrichter H, Trautmann M, Wischer B, Sommerauer S, Doppler B (2009) Unterrichten in heterogenen Gruppen: Das Qualitätspotenzial von Individualisierung, Differenzierung und Klassenschülerzahl Bd. 2. Nationaler Bildungsbericht Österreich, S 341–360 <https://doi.org/10.17888/nbb2009-2>
- Bayly-Castaneda K, Ramirez-Montoya MS, Morita-Alexander A (2024) Crafting personalized learning paths with AI for lifelong learning: a systematic literature review. *Front Educ* 9:1424386. <https://doi.org/10.3389/feuc.2024.1424386>
- Caines A et al (2023) On the application of large language models for language teaching and assessment technology. *arXiv preprint arXiv:2307.08393*
- Chen L, Chen P, Lin Z (2020) Artificial intelligence in education: a review. In: *IEEE Access*, S 75264–75278 <https://doi.org/10.1109/ACCESS.2020.2988510>

- Davis FD (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MISQ* 13(3):319–340. <https://doi.org/10.2307/249008>
- Desai D (2020) Modeling personalized E-learning for effective distance education. *Int J Res Appl Sci Eng Technol* 8:2428–2435. <https://doi.org/10.22214/ijraset.2020.6390>
- Gobara S, Kamigaito H, Watanabe T (2024) Do LLMs implicitly determine the suitable text difficulty for users? arXiv preprint arXiv:2402.14453
- Hein L et al (2024) ChatGPT als Unterstützung von Lehrkräften-Einordnung, Analyse und Anwendungsbeispiele. *HMD* 61:449–470. <https://doi.org/10.1365/s40702-024-01052-9>
- Holstein K, Alevan V (2022) Designing for human-AI complementarity in K-12 education. *Ai Mag* 43(2):239–248. <https://doi.org/10.1002/aaai.12058>
- Huang MH, Rust RT (2024) The caring machine: feeling AI for customer care. *J Mark*. <https://doi.org/10.1177/00222429231224748>
- Ifelebuegu AO (2023) Rethinking online assessment strategies: authenticity versus AI chatbot intervention. *J Appl Learn Teach* 6:385–392. <https://doi.org/10.37074/jalt.2023.6.2.2>
- Jin H, Kim Y, Park YS, Tilekbay B, Son J, Kim J (2024) Using large language models to diagnose math problem-solving skills at scale. In: Proceedings of the eleventh ACM conference on learning@ scale, S 471–475 <https://doi.org/10.1145/3657604.3664697>
- Karademir Coskun T, Alper A (2024) Evaluating the evaluators: a comparative study of AI and teacher assessments in higher education. *Digit Educ Rev* 45:124–140. <https://doi.org/10.1344/der.2024.45.124-140>
- Karpenko OM et al (2019) Individualization of learning: an investigation on educational technologies. *J Hist Cult Art Res* 8(3):81–90. <https://doi.org/10.7596/taksad.v8i3.2243>
- Kasneci E et al (2023) ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individ Differ* 103:102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Klafki W, Stöcker H (1976) Innere Differenzierung des Unterrichts. *Z Pädagog* 22(4):497–523
- Kleinberg J et al (2018) Discrimination in the age of algorithms. *J Leg Anal* 10:113–174. <https://doi.org/10.1093/jla/laz001>
- Leiker D, Finnigan S, Gyllen AR, Cukurova M (2023) Prototyping the use of Large Language Models (LLMs) for adult learning content creation at scale. arXiv preprint arXiv:2306.01815
- Lewis P et al (2020) Retrieval-augmented generation for knowledge-intensive NLP tasks. *Adv Neural Inf Process Syst* 33:9459–9474. <https://doi.org/10.48550/arXiv.2005.11401>
- Li L, Jiang Z (2013) A professional learning of the implementation of individualized education in cultivating mode of innovative talents—A case study of individualized education in Hebei university of science and technology. In: Proceedings of the International Conference on Education Technology and Information System (ICETIS 2013). Atlantis Press, S 606–609 <https://doi.org/10.2991/icetis-13.2013.135>
- MacQueen JB (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Bd. 1, S 281–297
- McDonald Connor C et al (2009) Individualizing student instruction precisely: effects of childx instruction interactions on first graders' literacy development. *Child Dev* 80(1):77–100. <https://doi.org/10.1111/j.1467-8624.2008.01247.x>
- Minn S (2022) AI-assisted knowledge assessment techniques for adaptive learning environments. *Comput Educ Artif Intell* 3:100050. <https://doi.org/10.1016/j.caeai.2022.100050>
- Nguyen H-H, Nguyen VA (2023) Personalized learning in the online learning from 2011 to 2021: a bibliometric analysis. *Int J Inf Educ Technol* 13(8):1261–1272. <https://doi.org/10.18178/ijiet.2023.13.8.1928>
- Noble SU (2018) Algorithms of oppression: how search engines reinforce racism. In: Algorithms of oppression. New York University Press, <https://doi.org/10.2307/j.ctt1pw9w5>
- Raggl A (2023) Praktiken des Helfens im individualisierten jahrgangsgemischten Unterricht. In: Schnitzer A, al (Hrsg) Schulische Praktiken unter Beobachtung. Springer VS, Wiesbaden, S 19–34 https://doi.org/10.1007/978-3-658-41952-3_2
- Rocein D, Curry AC, Hovy D (2023) Know your audience: do LLMs adapt to different age and education levels? arXiv preprint arXiv:2312.02065
- Ruiz JG, Mintzer MJ, Leipzig RM (2006) The impact of e-learning in medical education. *Acad Med* 81(3):207–212. <https://doi.org/10.1097/00001888-200603000-00002>
- Sandberg G (2022) The use of digital technology for differentiation of teaching in early school years. *Nord J Lit Res* 8(1):172–188. <https://doi.org/10.23865/njlr.v8.3104>

- Schaufeli WB, Martinez IM, Pinto AM, Salanova M, Bakker AB (2002) Burnout and engagement in university students: a cross-national study. *J Cross Cult Psychol* 33(5):464–481. <https://doi.org/10.1177/0022022102033005>
- Schmid U, Blanc B, Toepel M (2021) KI@Bildung. Lehren und Lernen in der Schule mit Werkzeugen Künstlicher Intelligenz (Schlussbericht im Auftrag der Deutsche Telekom Stiftung. Bonn/Berlin/ Essen 2021)
- Shihab SR et al (2023) Educational technology in teaching community: reviewing the dimension of integrating ed-tech tools and ideas in classrooms. *Eduvest J Univers Stud* 3(6):1028–1039. <https://doi.org/10.59188/eduvest.v3i6.835>
- St-Hilaire F et al (2022) A new era: intelligent tutoring systems will transform online learning for millions <https://doi.org/10.48550/arXiv.2203.03724>
- Streber D et al (2015) Grundwissen Lehrerbildung: Umgang mit Heterogenität – Praxisorientierung, Fallbeispiele, Reflexionsaufgaben. Cornelsen
- Tetzlaff L, Schmiedek F, Brod G (2021) Developing personalized education: a dynamic framework. *Educ Psychol Rev* 33(3):863–882. <https://doi.org/10.1007/s10648-020-09570-w>
- Tomlinson CA (2001) How to differentiate instruction in mixed-ability classrooms. ASCD
- Tukey JW (1949) Comparing individual means in the analysis of variance. *Biometrics* 5(2):99–114. <https://doi.org/10.2307/3001913>
- Vaganova OI et al (2020) The role of educational technologies in the individualization of educational students' activities. *Amaz Investig* 9(28):552–559. <https://doi.org/10.34069/AI/2020.28.04.60>
- Vaswani A et al (2017) Attention is all you need. *Adv Neural Inf Process Syst*. <https://doi.org/10.48550/arXiv.1706.03762>
- Vereinte Nationen 17 Ziele – Ziele für Nachhaltige Entwicklung – Agenda 2030 der UN. <https://17ziele.de/ziele/4.html>. Zugegriffen: 24. Mai 2024

Hinweis des Verlags Der Verlag bleibt in Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutsadressen neutral.