



Dem geschichtsspezifischen Einsatz von KI auf der Spur

Eine hochschuldidaktische Untersuchung der Anwendungsstrategien von Lehramtsstudierenden im Kontext von Geschichtsunterricht

ISABEL ELSNER & FRANZISKA PILZ

Zusammenfassung

Künstliche Intelligenz hält spätestens seit der Veröffentlichung von ChatGPT weltweit Einzug in Klassenzimmer und stellt (zukünftige) Geschichtslehrkräfte vor neue Herausforderungen. Die Geschichtsdidaktik als verantwortliche Disziplin ist gefordert, diesen mit theoretischen Reflexionen zum historischen Lernen im Kontext der Einflüsse von KI sowie geschichtsdidaktischen Impulsen zur Unterstützung der Lehrkräfte in der Praxis zu begegnen. Kontrovers verhandelt wird derzeit, worin überhaupt das „Fachspezifische“ am KI-gestützten historischen Lernen liegt bzw. liegen kann. Die vorliegende Studie soll einen Beitrag zu dieser Diskussion leisten, indem sie diese um die Perspektive der Lehramtsstudierenden erweitert. In einem Pre-Post-Setting konnten geschichtsdidaktische Anwendungsstrategien zum Einsatz von KI im Geschichtsunterricht erhoben werden, die zeigen, wie die zukünftigen Lehrkräfte z. B. mit ChatGPT historisches Lernen im Geschichtsunterricht planen und welche Schwerpunkte sie dabei setzen.

Schlüsselwörter: Künstliche Intelligenz; ChatGPT; Geschichtsdidaktik; Geschichtsunterricht; Hochschullehre

On the Track of the History-Specific Use of AI

A University Didactic Study of Student's Application Strategies in the Context of History Teaching

Abstract

Since the publication of ChatGPT at the latest, artificial intelligence has found its way into history classrooms worldwide and is presenting (future) history teachers with new challenges. History didactics can help meeting those challenges by offering theoretical reflections on historical learning with AI as well as providing impulses to support teachers in practice. The subject-specific about AI-supported historical learning is currently debated. This study aims to contribute to this debate by includ-

ing student's perspectives. Therefore, a pre-post setting for collecting the data has been chosen. It gives insights into student's application strategies and shows how future teachers would include AI into their history lessons and which priorities they set.

Keywords: Artificial Intelligence; ChatGPT; History Didactics; History Lessons; Academic Teaching

1 Einleitung

„Als Teil des [Geschichts-] Unterrichts ist KI nicht geeignet. Es ist sogar, nahezu in jeder Hinsicht, extrem gefährlich“¹, so eine studentische Perspektive auf die Frage, wie KI im Geschichtsunterricht eingesetzt werden könnte. „Wichtig ist mir, dass die SuS den kritischen Umgang mit KI erlernen und vor allem differenzieren lernen, welche Vor- aber auch Nachteile sie hat. [...] Grenzen und Möglichkeiten sollten dabei aufgezeigt werden, sodass die SuS bewusst mit KI umgehen können“², wurde im gleichen Kontext von einem Kommilitonen angeführt.

Diese Überlegungen von Lehramtsstudierenden des Fachs Geschichte im Bachelor, die gerade ihr geschichtsdidaktisches Proseminar absolvierten, geben einen Einblick in die kontroversen Diskussionen, die zum Thema *Künstliche Intelligenz* (KI) gesellschaftlich geführt werden. Sie stehen auch exemplarisch für Positionierungen, die innerhalb des geschichtsdidaktischen Diskurses in Erscheinung treten. Doch zwischen strikter Ablehnung und konstruktivem Umgang wird vor allem die Relevanz geschichtsdidaktischer Reflexionen rund um den (möglichen) Einsatz von KI im Geschichtsunterricht deutlich. Adressiert werden verständlicherweise die Fachdidaktiken, die (bezogen auf das Fach Geschichte) in Teilen bereits reflektierte Anregungen zu Umgangsmöglichkeiten mit generativen Sprachmodellen im Geschichtsunterricht (z. B. Held, 2023; Neubert & Burkhardt, 2024) vorweisen können.³ Kontrovers verhandelt wird innerhalb des geschichtsdidaktischen Diskurses vor allem die Frage nach dem Fachspezifischen.⁴ Von der Antwort auf diese Frage scheinen auf eine grundsätzliche Art und Weise die Legitimität *oder* die Notwendigkeit *oder* gar die Gebotenheit des Einsatzes generativer Sprachmodelle im Geschichtsunterricht sowie in der Ausbildung der Geschichtslehrkräfte abzuhängen. Eine Antwort auf diese Frage kann hier nicht erfolgen. Doch den eingangs angeführten Äußerungen der Bachelorstudierenden wohnt eine hochgradig aktuelle geschichtsdidaktische Brisanz inne. Obwohl mit ihnen die Frage nach dem Geschichtsspezifischen nicht beantwortet werden kann, erweitern sie den Diskurs um eine wichtige Facette: die studentische Perspektive.

Da spätestens seit der Einführung von ChatGPT⁵ als kostenloses Werkzeug zu KI-gestützter Texterzeugung im November 2022 KI zum Bestandteil der Lebenswelt vieler Schülerinnen und Schüler geworden ist, beeinflusst sie auch das Lernen und Lehren im Hochschulkontext. Sowohl Lernende als auch Lehrende benötigen umfassende Kompetenzen, um einen lernförderlichen Umgang mit Large-Language-Models (LLM) erfahren und aktiv gestalten zu können, so die Ständige Wissenschaftliche Kommission der Kulturministerkonferenz (SWK) (vgl. SWK, 2024, S. 19). Daraus lassen sich neben fachdidaktischen Implikationen grundlegende Konsequenzen für das akademische Lehren und Lernen ableiten. Zum einen sollten zukünftige Geschichtslehrkräfte im Studium zu einem reflektierten Umgang mit LLM befähigt werden, um Potenziale und Grenzen textgenerierender KI im Kontext von Geschichtsunterricht einschätzen und reflektieren zu können. Zudem ist die Präsenz von KI im Schulkontext nur noch bedingt optional: Die SWK empfiehlt, den regelmäßigen Einsatz von LLM als Schreibunterstützung ab der achten Jahrgangsstufe zu erwägen; ab der

1 Zitat aus der ersten Erhebung, Fall SD30Ö01

2 Zitat aus der ersten Erhebung, Fall SP30P03

3 Alexandra Krebs stellt in ihrer Dissertation mit der „App in die Geschichte“ ein Beispiel vor, wie KI im Geschichtsunterricht genutzt werden kann, um historisches Lernen zu fördern. Krebs 2024.

4 Dabei kann beobachtet werden, dass, während zunehmend konstruktive Auseinandersetzungen veröffentlicht werden, gleichzeitig kritische Stimmen, die dem grundsätzlichen Einsatz von KI im Kontext geschichtsdidaktischen Lehrens und Lernens eher ablehnend gegenüberstehen, sich vor allem auf Konferenzen Gehör verschaffen. Die real existierenden Positionierungen finden zumindest in dieser Deutlichkeit selten Eingang in Publikationen.

5 ChatGPT steht exemplarisch für ein Large-Language-Model (LLM) im Kontext der Studie.

Sekundarstufe II wird ein vielfältiger Einsatz zur Erstellung von Texten verschiedener Genres befürwortet (vgl. SWK, 2024, S. 20).

Zum anderen sollten Studierende Ideen entwickeln können, inwiefern ChatGPT im Kontext von Geschichtsunterricht sinnvoll eingesetzt werden kann; von der Unterrichtsvorbereitung bis zur -durchführung. Dafür lassen sich noch weitere Gründe anführen: Textgenerierende KI kann, entsprechend verwendet, Lehrkräfte entlasten und mehr Kapazitäten für den eigentlichen Unterricht schaffen. Zudem konnte John Hattie zeigen, dass Chatbots das Potenzial haben, bemerkenswerte Effekte (Effektstärke 0,67) auf die Lernleistung von Kindern und Jugendlichen zu erzielen (Hattie, 2008; Zierer, 2024).⁶ Dieser Effekt tritt nur dann auf, wenn der Einsatz von Chatbots in eine didaktische Planung und Begleitung eingebettet ist (Zierer, 2024, S. 38). Die Wirksamkeit ist damit an die Unterrichtsqualität gebunden (Zierer, 2024, S. 39). Als Tool in den Geschichtsunterricht implementiert, kann ChatGPT auf ganz unterschiedliche Weise historisches Lernen befördern, so eine geschichts-didaktische Perspektive (Held, 2024). Da davon auszugehen ist, dass Schülerinnen und Schüler es unabhängig davon nutzen, ob der Einsatz sinnvoll oder erwünscht ist, sollten die Studierenden in die Lage versetzt werden, ihre zukünftigen Lerngruppen im Umgang mit LLM anzuleiten. Daher sollten sie die Möglichkeit erhalten, über Chancen und Grenzen von ChatGPT zu diskutieren. Nicht zuletzt könn(t)en Kompetenzen im Umgang mit LLM auch historisches Lernen während des Studiums fördern und das Studium bereichern.

Die Ausgangslage für die vorliegende Studie speist sich aus mehreren Einflussfaktoren: Erstens wird eine fachdidaktische Diskussion darüber geführt, inwiefern KI-gestütztes Lernen historisches Lernen sein kann. Zweitens sind Studierende wie Lehrende, wir selbst also, mit der Situation konfrontiert, dass KI in Form unterschiedlichster Anwendungen längst Eingang in schulisches wie universitäres Lehren und Lernen gefunden hat. Aus dieser Erkenntnis entwickelte sich unser Bedürfnis, das Thema KI in der geschichts-didaktischen Lehre proaktiv als Forscherinnen anzugehen. Daraus resultierte, dass wir in unserem Projekt jeweils in einer Doppelrolle fungieren: erstens in der Rolle der Lehrenden, zweitens als Forscherin. Die Forschungsfrage soll diesen Einflussfaktoren Rechnung tragen: Inwiefern zeigen sich fachspezifische Anwendungsstrategien zum Einsatz von KI im Kontext von Geschichtsunterricht bei Lehramtsstudierenden des Fachs Geschichte?

Dies wurde im Sommersemester 2024 im Rahmen der gleichnamigen Proseminare „Geschichts-didaktik im schulischen Kontext am Beispiel von historischen Kontroversen im KI-gestützten Geschichtsunterricht“ erforscht, genauer gesagt im seminarbegleitenden Forschungsprojekt. Die Seminare boten den Studierenden eine Anwendungs- und Diskussionsplattform, die es ermöglichte zu beobachten, wie Studierende im Umgang mit KI Konzepte historischen Lernens mindestens fach-didaktisch zu wenden versuchten. Hierin liegt möglicherweise auch eine Chance, dem Fachspezifischen auf die Spur zu kommen...

2 Konzeption des Projekts

2.1 Einbettung des Forschungsprojekts in die geschichts-didaktische Hochschullehre

Das Forschungsprojekt wurde im Sommersemester 2024 im Rahmen zweier geschichts-didaktischer Proseminare durchgeführt. Der Titel der Seminare lautete: „Geschichts-didaktik im schulischen Kontext am Beispiel von historischen Kontroversen im KI-gestützten Geschichtsunterricht“. Die Studierenden, die an der Studie teilnahmen, befanden sich zum Zeitpunkt der Erhebungen i. d. R. zwischen dem dritten und dem sechsten Semester im *Bachelor of Education* (Lehramt, Geschichte, GyGe/HRSG). Die Seminare waren als praxisbezogenes Kollaborationsprojekt angelegt. In einer ersten Seminarphase setzten sich die Studierenden mit geschichts-didaktischer Theorie auseinander und diskutierten vor diesem Hintergrund verschiedene Einsatzmöglichkeiten von KI im Geschichtsunterricht. In der daran anschließenden Praxisphase setzten sie sich praxisorientiert mit dem Einsatz von

6 Hattie stützt sich hier auf Metaanalysen: Lee & Hwang, 2022; Alemdag, 2023; Wu et al., 2023; Deng & Yu, 2023; Zhang et al., 2023.

KI im Kontext von Geschichtsunterricht auseinander. Ein Schwerpunkt in dieser Phase war die Vorbereitung und Planung eines Besuches im MINT-Schülerlabor der Universität Paderborn. Dort erhielten die Studierenden die Möglichkeit, ihre Unterrichtsplanungen mit Schülerinnen und Schülern einer elften Klasse zu verschiedenen historischen Themen umzusetzen. Der Schwerpunkt lag dabei auf dem Einsatz von KI. Abschließend fand eine Reflexionsphase statt, in der KI u. a. als Planungs- sowie Arbeitstool im Kontext von Geschichtsunterricht reflektiert wurde. Die Frage nach den Chancen und Grenzen für historisches Lernen stand dabei im Mittelpunkt.



Abbildung 1: Aufbau der Proseminare

Im Proseminar sollten die Studierenden über das gesamte Semester zentrale geschichtsdidaktische Ansätze sowie die erkenntnistheoretischen Grundlagen kennen- und bewerten lernen. Zudem sollten sie digitale Medien und Methoden des Geschichtslernens und -unterrichts sowie deren fachspezifische Möglichkeiten und Grenzen kennen- und reflektieren lernen (Präsidium der Universität Paderborn, 31.05.2022, S. 18). Das Setting, in dem die Studierenden gelernt haben, war dabei grundlegend innovativ: Erstens bestand das gesamte Semester über eine Kooperation mit dem jeweils anderen geschichtsdidaktischen Proseminar. Auf diese Weise profitierten die Studierenden beider Seminare von unvoreingenommenem Peer-Feedback. Zweitens planten beide Seminare im letzten Drittel des Semesters gemeinsam Geschichtsunterrichtssequenzen, die die Studierenden in einer Blocksitzung im Schülerlabor der Universität Paderborn mit einer Schulklasse umsetzten.

Für uns bedeutete die Einbettung des Forschungsprojekts in unsere Lehre von Beginn an eine Reflexion unserer Doppelrolle. Auch das Verhältnis von Lehre und Forschung innerhalb des Projekts musste reflektiert werden. Die Lehre sollte im Sinne der teilnehmenden Studierenden der Forschung übergeordnet sein. So überlegten wir, wie Lehr- und Forschungsprojekt voneinander profitieren können. Das Ziel sollte eine symbiotische Verbindung der Lehrveranstaltung mit dem Forschungsprojekt sein.

2.2 Erkenntnisziel und Fragestellung

Dem Aufsatz liegt die Annahme zugrunde, dass der professionelle Umgang mit KI im Rahmen von (Geschichts-) Unterricht den Einsatz von KI in der (geschichtsdidaktischen) Hochschullehre voraussetzt. Dementsprechend war auch das Ziel dieses Seminarprojekts ein doppeltes: Einerseits standen fachspezifische Interessen hinsichtlich des Anwendungsbezugs von KI im Geschichtsunterricht im Vordergrund. Andererseits ging es um eine fachunspezifische Auseinandersetzung mit den Einsatzmöglichkeiten von KI in der Hochschullehre.

Das Forschungsprojekt wiederum legt den Fokus auf die Erweiterung des geschichtsdidaktischen Diskurses um den fachspezifischen Einsatz von KI im Kontext von Geschichtsunterricht durch die Ergänzung studentischer Perspektiven. Die Studie fragt danach, inwiefern sich fachspezifische Anwendungsstrategien zum Einsatz von KI im Kontext von Geschichtsunterricht bei Lehramtsstudierenden des Fachs Geschichte zeigen. Diese können sich sowohl auf den Einsatz von KI als Planungstool als auch auf die Durchführung von Geschichtsunterricht beziehen.

Im Folgenden sollen zunächst theoretische und methodische Prämissen vorgestellt werden, bevor die Ergebnisse der Studie diskutiert werden.

3 Theorie: KI im Kontext von Hochschullehre und Geschichtsunterricht

Um dem leitenden Erkenntnisinteresse nachzugehen, ist eine theoretische Rahmung auf zwei Ebenen erforderlich: Erstens müssen geschichtsdidaktische Überlegungen und Befunde hinsichtlich historischen Lehrens und Lernens berücksichtigt werden, um die Umgangsweisen der Studierenden einordnen und fachspezifische Anwendungsstrategien identifizieren zu können. Zweitens ist die hochschuldidaktische Perspektive relevant, die eine vergleichsweise breite empirische Basis zur Einordnung der Befunde offeriert.

3.1 KI in der Geschichtsdidaktik

Um herausarbeiten zu können, inwiefern die untersuchten Lehramtsstudierenden fachspezifische Anwendungsstrategien beim Einsatz von KI zeigen, müssen solche zunächst definiert werden. Dies soll im Folgenden geschehen. Im geschichtsdidaktischen Diskurs werden zunehmend theoretische Überlegungen und unterrichtspragmatische Vorschläge vorgelegt, die KI-gestütztes historisches Lernen zum Gegenstand haben. Zwecks Einordnung und Anbindung dieser Studie an den Fachdiskurs werden sie angeführt. Differenziert werden muss hier zwischen Publikationen zum Geschichtsunterricht und geschichtsdidaktischer Lehre an Hochschulen. Vorweggenommen werden kann hier, dass bisher kaum empirische Befunde zum Einsatz von KI in geschichtsdidaktischen Lehr-Lernkontexten publiziert wurden.⁷

Das Vorliegen einer fachspezifischen Anwendungsstrategie wird angenommen, wenn die Studierenden in ihre Überlegungen zum Einsatz von KI im Geschichtsunterricht Modelle oder Konzepte historischen Denkens und Lernens einbezogen haben. Dabei konnte es sich etwa auf konzeptioneller Ebene um die grundständige Berücksichtigung von Theorie handeln, die expliziert wurde, beispielsweise im Kontext der Frage danach, was Geschichte ist (z. B. Rüsen, 2003), oder hinsichtlich der Annahme des Vorhandenseins von Geschichtsbewusstsein sowie der Förderung eines reflektierten Geschichtsbewusstseins (z. B. Schönemann, 2012; Schreiber, 2008).⁸ Insbesondere „Historisches Lernen“ fand in vielen Fällen Erwähnung (Meyer-Hamme, 2018). In diesem Bereich ist auch die Auseinandersetzung mit oder Berücksichtigung von Narrativität zu verorten (Barricelli, 2011). Zudem wurden auch unterrichtspragmatische Überlegungen angestellt, die geschichtsdidaktische Prinzipien

⁷ Die Dissertation von Alexandra Krebs bietet eine erste Erkundung dieses Forschungsfeldes. Krebs 2024.

⁸ Diese Themen stehen exemplarisch für viele weitere Modelle und Konzepte, die von den Studierenden bedacht werden könnten. Es soll und kann keinesfalls Anspruch auf Vollständigkeit erhoben werden. An dieser Stelle kann nicht geleistet werden, die einzelnen Themenbereiche auszuführen und zu erklären, weil dies den Rahmen dieses Aufsatzes überschreiten würde. Die angeführte Auswahl bezieht sich auf die Literatur, mit der sich die Studierenden im Laufe des Semesters bzw. im Vorfeld im Kontext der vorgeschalteten Vorlesung befasst haben.

wie Problemorientierung (Barricelli, 2016), Multiperspektivität (Bergmann, 2016) oder Handlungsorientierung (Völkel, 2016) berücksichtigten. Auch Überlegungen zu Lebens- und Gegenwartsbezügen (Bergmann, 2016), Quellen- (Pandel, 2013) und Wissenschaftsorientierung (Borries, 2013) wurden einbezogen.

Zum Einsatz von KI im Kontext von Geschichtsunterricht liegen mehrere Publikationen vor.

Mayer führt didaktische Überlegungen zu einer Unterrichtsidee aus, in der der Fokus auf der Entwicklung und Schulung der Fragekompetenz liegt (Mayer, 2023). Zudem sollen Kompetenzen im Bereich des kritischen Denkens, des Argumentierens sowie des Interpretierens geschult werden (Mayer, 2023). Die Chats sollen der Lehrkraft anschließend u. a. zeigen können, „[...] wie gut die Lernenden in der Lage sind, die Aussagen von ‚historischen Persönlichkeiten‘ kritisch zu hinterfragen und zu bewerten, insbesondere im Hinblick auf ihre eigenen Werte und Überzeugungen“ (Mayer, 2023).

Wegener diskutiert Chancen und Herausforderungen des Einsatzes von ChatGPT im Geschichtsunterricht (Wegener, 2023). Er erwägt einen reflektierten und kritischen Umgang mit ChatGPT als Ziel von Geschichtsunterricht. Seine Eignung dafür liege in der quellenkritischen Methodik begründet (Wegener, 2023).

Held wiederum legt eine Sammlung an Unterrichtsvorschlägen vor, in der anhand unterschiedlicher historischer Sachverhalte auf verschiedene Weise mit ChatGPT historisch gelernt werden soll (Held, 2024). Damit verbunden schlägt er eine Systematisierung des KI-Einsatzes im Unterricht vor. KI kann im Geschichtsunterricht demnach als Mittel zur (1) Generierung von Materialien, zur (2) Analyse von Materialien oder zur (3) Simulation von Interaktionen genutzt werden (Held, 2024). Lernende sollten zu einem verantwortungsvollen Umgang mit KI befähigt werden, die u. a. Quellen und Darstellungen analysieren könnte (Held, 2024). Die Charakterisierung von KI als ein Hilfsmittel müsse dabei zentral sein (Held, 2024).

Burkhardt & Neubert stellen ausgehend von einem Bedarf an geschichtsdidaktischen Impulsen seitens der Lehrkräfte mit Anregungen verbundene Überlegungen zum Umgang mit generativen Sprachmodellen wie ChatGPT im Geschichtsunterricht an (Burkhardt & Neubert, 2024). Ihre Praxisvorschläge fokussieren Methodenkompetenz, historische Bildquellen und Fragekompetenz (in Verbindung mit Prompt Engineering). Sie zeigen u. a., inwiefern die Rüsen'schen Kriterien der empirischen, normativen und narrativen Triftigkeit (Rüsen, 1983) auch als Kriterien für die Analyse KI-generierter Narrationen nutzbar sind (Burkhardt & Neubert, 2024). Sie weisen deutlich auf die Notwendigkeit hin, „KI-generierte Geschichte(n) an Quellen rückzubinden, zu dekonstruieren, historische [...] Triftigkeiten zu prüfen und die geschichtskulturelle Relevanz zu thematisieren“ (Burkhardt & Neubert, 2024). Postuliert werden die Einbindung des Themas „KI“ in die universitäre Ausbildung der Geschichtslehrkräfte sowie vom Verständnis historischen Lernens mit KI als Ko-Konstruktion einer Mensch-Maschine-Interaktion auszugehen und dies als Basis für theoretische geschichtsdidaktische sowie für methodische Überlegungen des Geschichtsunterrichts zu betrachten (Burkhardt & Neubert, 2024).

3.2 KI in der Hochschullehre

Das hochschulpolitische Interesse am Einsatz KI-basierter Tools in Lehr-Lernkontexten wächst in Deutschland nach Jahrzehnten der zurückhaltenden Auseinandersetzung mit der Thematik erst seit Kurzem und umso rasanter (Schmohl et al., 2023, 8). Der mögliche Einsatz von KI in der Hochschulbildung, im internationalen Kontext oft als „AI in Education“ (AIED) bezeichnet, ist bereits seit ungefähr 30 Jahren Forschungsgegenstand verschiedener akademischer Disziplinen (Schmohl et al., 2023, 8). Die theoretische Auseinandersetzung mit dieser Thematik reicht noch weiter zurück (Schmohl et al., 2023, 8). Der Diskurs über KI in der Hochschulbildung nimmt parallel zur technischen Entwicklung Fahrt auf (Schmohl et al., 2023, 9).

KI-Anwendungen, die im Hochschulkontext eingesetzt werden, lassen sich grundsätzlich in drei Ebenen unterteilen (Ulrich & Heckmann, 2017; Watanabe, 2023): In Bezug auf den Einsatz im Kontext konkreter Lehr-Lernprozesse können KI-Anwendungen auf der Mikroebene zum Einsatz

kommen (Schmohl et al., 2023, 10). Die Integration von KI in die Gestaltung des Curriculums geschieht auf der Mesoebene, während auf der Makroebene die Chancen des Einsatzes von KI auf der institutionellen Ebene der Hochschule zu verorten sind (Schmohl et al., 2023, 11). Relevant ist in diesem Kontext insbesondere die Mikroebene, die Lehr-Lernprozesse fokussiert. Grundsätzlich kann KI entweder als Thema oder als Tool auf dieser Ebene betrachtet werden (Watanabe, 2023; Wannemacher & Bodmann, 2021).

Im Rahmen der vorliegenden empirischen Studie stand, bezogen auf das Seminarsetting, Lehre mit KI im Fokus. Im Sinne der erkenntnisleitenden Fragestellung soll untersucht werden, inwiefern fachspezifische Anwendungsstrategien identifiziert werden können. An dieser Stelle sollen solche empirischen Studien herangezogen werden, die es ermöglichen, den Umgang der Studierenden mit KI einordnen zu können.

Eine erste deutschlandweite Untersuchung des Einsatzes und der Nutzung von KI-Systemen durch Studierende im Kontext von Studium und Lehre fand 2023 als Online-Befragung von Studierenden statt, die zum Befragungszeitpunkt an einer deutschen Hochschule/Universität eingeschrieben waren (Garrel et al., 2023). In der Studie konnte gezeigt werden, dass 49 % aller befragten Studierenden (N=6311) ChatGPT/GPT-4 im Kontext ihres Studiums eingesetzt haben (Garrel et al., 2023, 26). Zudem wurde dabei erhoben, wofür im Rahmen des Studiums KI genau eingesetzt wurde, z. B. für Recherchen (Garrel et al., 2023, 27). Ein relevantes Ergebnis für dieses empirische Vorhaben ist, dass Garrel et al. gezeigt haben, dass viele Studierende insbesondere an dem „wissenschaftlichen Output“ interessiert sind, den sie mittels KI zu generieren glauben. Wie dieser jedoch zustande kommt, scheint für die meisten Studierenden eher zweitrangig. An dieser Stelle sehen die Forschenden die Hochschulen in der Verantwortung, den Studierenden den Umgang mit KI als Instrument zu vermitteln und dabei auch Grenzen bei der Nutzung aufzuzeigen (Garrel et al., 2023, 38). Die Studie ist relevant, weil sie viele Informationen über das KI-Nutzungsverhalten von Studierenden enthält.

In einer Kurzstudie konnte Bhattacharya zeigen, dass 34,9 % der Befragten (N=1205) KI-Bots wie ChatGPT zum Lernen bzw. für das Studium nutzten (Bhattacharya, 2023, 2).

Ferner interessant ist eine Studie, in der fachspezifisches Wissen von Lehrkräften zu KI erhoben wurde (Lindner & Romeike, 2019). Relevant ist sie, weil das Wissen über KI fokussiert wurde. Das Ergebnis ist mit Vorsicht auf Studierende zu beziehen, weil es sich bei den Studienteilnehmenden um fertig ausgebildete Lehrkräfte handelte, denen Schlagworte präsentiert wurden, deren Bezug zu KI diese einordnen sollten. Das Wissen der Lehrkräfte über Anwendungen zu KI wird als zumeist defizitär herausgestellt (Lindner & Romeike, 2019, 5 ff.). Dies muss nicht zwingend auf die Studierenden zutreffen.

4 Methodische Überlegungen

4.1 Datenerhebung

Um die o. g. Fragestellung(en) zu beantworten, wurde ein Pre-/Posttest-Design für die Studie gewählt.⁹ Die Erhebungen fanden jeweils in Form schriftlicher Befragungen statt. Zu Beginn des Semesters wurde in der ersten Seminarsitzung die erste Erhebung durchgeführt. Die zweite Erhebung fand am Ende des Semesters in der Reflexionssitzung (siehe Kapitel 2) statt.¹⁰ In Form der schriftlichen Befragung ließen sich die Erhebungen gut in die Seminarsitzungen integrieren und konnten von den Studierenden zugleich zur Selbstreflexion genutzt werden. Vor Beginn der Erhebung wurden die Studierenden über das Ziel und die Konzeption der Studie informiert. Die Teil-

⁹ Trotz des ähnlichen Forschungsdesigns handelt es sich bei der Studie nicht um eine Interventionsstudie, da nicht die Wirksamkeit des Seminars, sondern die Entwicklung der Studierenden im Vordergrund stand. Eine Interventionsstudie erschien zudem nicht praktikabel, weil dafür keine Vergleichsgruppe zur Verfügung stand. Zudem waren die Einflussfaktoren auf die Entwicklung der Studierenden sehr vielfältig und nicht allein auf das Seminar beschränkt, wie die Ergebnisse zeigen.

¹⁰ Die zweite Befragung wurde zu Beginn der Reflexionssitzung durchgeführt. So konnte sichergestellt werden, dass ausschließlich individuelle Ansichten, nicht die Ergebnisse der Seminare Diskussion in die Antworten einfließen.

nahme an der Studie erfolgte freiwillig und unabhängig von sonstigen Seminarleistungen. Beide Befragungen fanden digital statt und wurden als anonyme schriftliche Befragung unter Beisein der Befragungsleitung durchgeführt (Aeppli et al., 2014, 174). Die Anonymität der Befragung sollte dazu beitragen, hierarchische Strukturen aufzuweichen und eine Zuordnung der Antworten zu den Seminarteilnehmenden zu verhindern, um so möglichst offene Antworten von den Studierenden zu erhalten.

Das Sample setzt sich aus insgesamt 26 Studierenden des Fachs Geschichte im *Bachelor of Education* zusammen. Zwar nahmen mehr als 30 Studierende an der Befragung teil, im Sample wurden jedoch nur die Abgaben derjenigen Studierenden berücksichtigt, die an beiden Erhebungen teilnahmen.¹¹

Die Befragungen setzen sich sowohl aus Multiple-Choice- als auch offenen Fragen zusammen und lassen sich somit als Mischform qualitativer und quantitativer Erhebungsformen charakterisieren. Die erste Erhebung zielte auf die Erhebung des (Vor-)Wissens der Studierenden über KI sowie den Einsatz von KI im Geschichtsunterricht ab. Die Anwendungsstrategien der Studierenden wurden dabei über die Frage erhoben, wie sie KI im Geschichtsunterricht einsetzen würden. Bei der zweiten Erhebung wurden dieselben Items abgefragt, um eine Vergleichbarkeit der Ergebnisse zu ermöglichen.

Die Studie lässt nur bedingt Aussagen über die Wirksamkeit der Seminare zu, da ein standardisiertes Vorgehen vor dem Hintergrund individueller Bedürfnisse der Seminargruppen nicht möglich war. Das hängt damit zusammen, dass die Qualität der Lehre auch vor dem Hintergrund der semesterbegleitenden Forschung stets oberste Priorität hatte. Zudem hatten auch andere Lehrveranstaltungen Einfluss auf die Antworten der Studierenden. Diese werden in der Erhebung zum Teil explizit benannt. Daher steht die Frage nach der Veränderung studentischer Anwendungsstrategien zu KI im Geschichtsunterricht im Vordergrund. Der Fokus liegt deswegen hauptsächlich auf den Antworten zu der offenen Fragestellung. Mithilfe der Multiple-Choice-Antworten lassen sich die Antworten der Studierenden jedoch kontextualisieren. Insofern sind die qualitativen Anteile der Erhebung den quantitativen Elementen übergeordnet. Aus diesem Grund erscheint es sinnvoll, auch bei der Auswertung der Daten qualitativ vorzugehen.

4.2 Datenauswertung

Die Daten wurden mithilfe der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz und Rädiker (2024) ausgewertet. Die zugrunde liegende Fragestellung wurde dabei bewusst offen formuliert, sodass die Befunde vorab nicht zu stark eingeschränkt wurden. Gleichzeitig war es wichtig, einen Anwendungsbezug in die Erhebung aufzunehmen und denselben Impuls für die erste und zweite Erhebung zu verwenden, um eine Vergleichbarkeit der Ergebnisse zu gewährleisten. Die anschließende Codierung erfolgte mithilfe von MAXQDA. Beim Codieren wurde das Verfahren des konsensuellen Codierens praktiziert, um die Güte der Codierungen überprüfen zu können (Hopf & Schmidt, 1993). Bei der Codierung der Studierendenantworten wurde sowohl deduktiv als auch induktiv vorgegangen. Die quantitativen Angaben konnten mithilfe deduktiver Kategorien codiert werden. Bei den qualitativen Antworten der Studierenden wurden hauptsächlich induktive Kategorien gebildet. Auf Grundlage der Fragestellung war zwar die Fachspezifik eine deduktive Kategorie, die übrigen Kategorien ergaben sich jedoch induktiv aus den Studierendenantworten.

Wie bereits angedeutet, wurde das Vorliegen einer fachspezifischen Anwendungsstrategie angenommen, wenn die Studierenden in ihre Überlegungen zum Einsatz von KI im Geschichtsunterricht Modelle oder Konzepte historischen Denkens und Lernens einbezogen haben. Die Qualität der studentischen Ausführungen wurde in diesem Kontext zunächst nicht bewertet. Sofern sich beide Forscherinnen darüber einig waren, dass ein fachspezifischer Bezug vorlag, wurde dieser als solcher codiert. Mit Blick auf die Forschungsfrage wurde dieser Code pro Fall ($n = \text{ein Proband}$) bei Vorlie-

¹¹ Über personalisierte Codes war eine Zuordnung beider Erhebungsteilnahmen möglich – eine personelle Zuordnung durch die Dozentinnen war hingegen nicht möglich.

gen einmal vergeben, damit eine Übersicht darüber gewonnen werden kann, wie viele der Teilnehmenden fachspezifische Überlegungen anstellen.

Um die Gültigkeit der Ergebnisse zu erhöhen, sind Zwischenergebnisse und Ergebnisse mit Geschichtsdidaktikerinnen und Geschichtsdidaktikern diskutiert worden, die nicht Teil des Projekts waren (Kuckartz & Rädiker, 2024). Eine Verallgemeinerbarkeit der Studienergebnisse ist nicht intendiert und steht in keinem Verhältnis zur Nützlichkeit der vorliegenden Erkenntnisse (Kuckartz & Rädiker, 2024).

5 Ergebnisse

Die Auswertung bestätigt die Relevanz der Fragestellung. Es kann gezeigt werden, dass 73 % der teilnehmenden Studierenden bereits vor Beginn des Seminars KI im Rahmen ihres Studiums nutzen. Die übrigen Studierenden geben an, grundsätzlich keine KI zu nutzen. Insgesamt haben die Teilnehmenden 23 verschiedene KI-Anwendungen bzw. konkrete Tools genannt.¹² Von 85 % der Studierenden wurde ChatGPT als ihnen bekanntes Beispiel für KI angeführt. Vor diesem Hintergrund erschien es für die Ausgestaltung des Seminarkonzeptes sinnvoll, ChatGPT als Beispielanwendung für ein LLM einzusetzen. Dieser Befund deckt sich mit den Ergebnissen von Garrel et al. (2023), die zeigen konnten, dass 49 % der von ihnen befragten Studierenden ChatGPT im Kontext ihres Studiums einsetzten (Garrel et al., 2023). Den Befunden von Bhattacharya (2023) widersprechen die Ergebnisse nicht.

Die Bereitschaft, KI in Zukunft als Lehrkraft im Geschichtsunterricht einzusetzen, liegt in der ersten Befragung bei 85 %. Nur ein Student spricht sich zu diesem Zeitpunkt gegen einen zukünftigen Einsatz von KI im Geschichtsunterricht aus, während drei Studierende angeben, unsicher in einer diesbezüglichen Entscheidung zu sein.

Im Kontrast dazu sinkt im Laufe des Seminars die Bereitschaft der potenziellen Lehrkräfte, KI im Geschichtsunterricht einzusetzen. In der zweiten Befragung geben nur noch 50 % der Studierenden an, KI zukünftig im Geschichtsunterricht einsetzen zu wollen. Die übrigen Teilnehmenden entscheiden sich dagegen oder sind sich in dieser Frage noch unsicher.

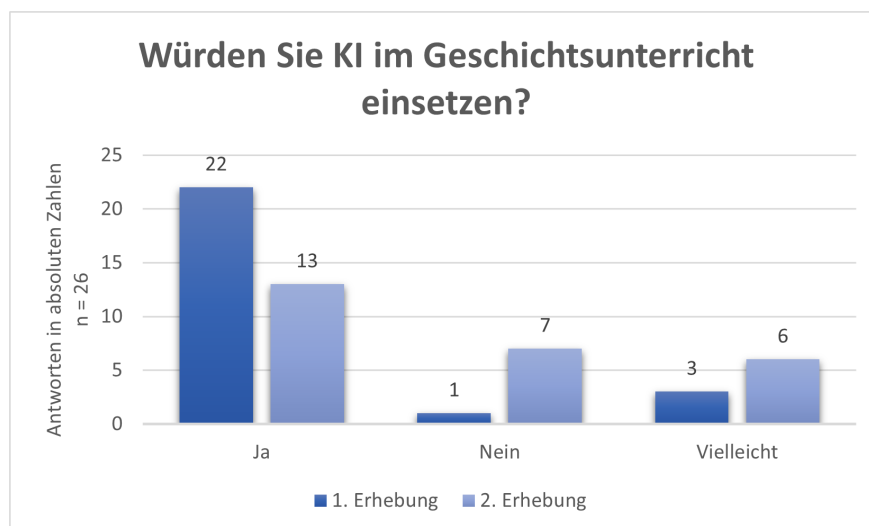


Abbildung 2: Bereitschaft der Studierenden (N=26), KI in Zukunft in von ihnen durchgeführtem Geschichtsunterricht einzusetzen. Ergebnisse aus der 1. und 2. Erhebung

¹² Folgende Beispiele für KI-Anwendungen haben die Studierenden genannt: ChatGPT (22), Perplexity AI (4), bildgenerierende KI (4), DeepL Translate (4), DALL-E (3), DeepL Write (3), Gemini Google AI (3), Stable Diffusion (2) Bard (1), ChatPDF (1), Smart Home (1), Sora AI (1), Charakter AI (1), Untertitelgenerierung bei Instagram (1), Finanz-Tools (1), Adobe Firefly (1), KI in Videospielen (1), KI in Computerspielen (1), (halb-) autonomes Fahren (1), Automatismen in der Robotik (1), Spracherkennungssoftware (1), Google Übersetzer (1), audio-generierende KI (1).

Zu differenzieren hiervon ist die Frage nach dem Einsatz von KI im Rahmen der Planung von Geschichtsunterricht. Dieser Befund ging aus der ersten Befragung hervor und wurde aus diesem Grund als Fragestellung in die zweite Befragung aufgenommen. Nach der Veranstaltung gaben 31 % der Befragten an, bei der Planung von Geschichtsunterricht KI einsetzen zu wollen, während sich 38 % dagegen aussprachen und weitere 31 % der Studierenden sich darüber noch nicht im Klaren waren.

Im Rahmen der offenen Frage, wie die Studierenden KI zukünftig konkret im Geschichtsunterricht einsetzen würden und was ihnen dabei wichtig wäre, ergaben sich mehrere induktive Kategorien.

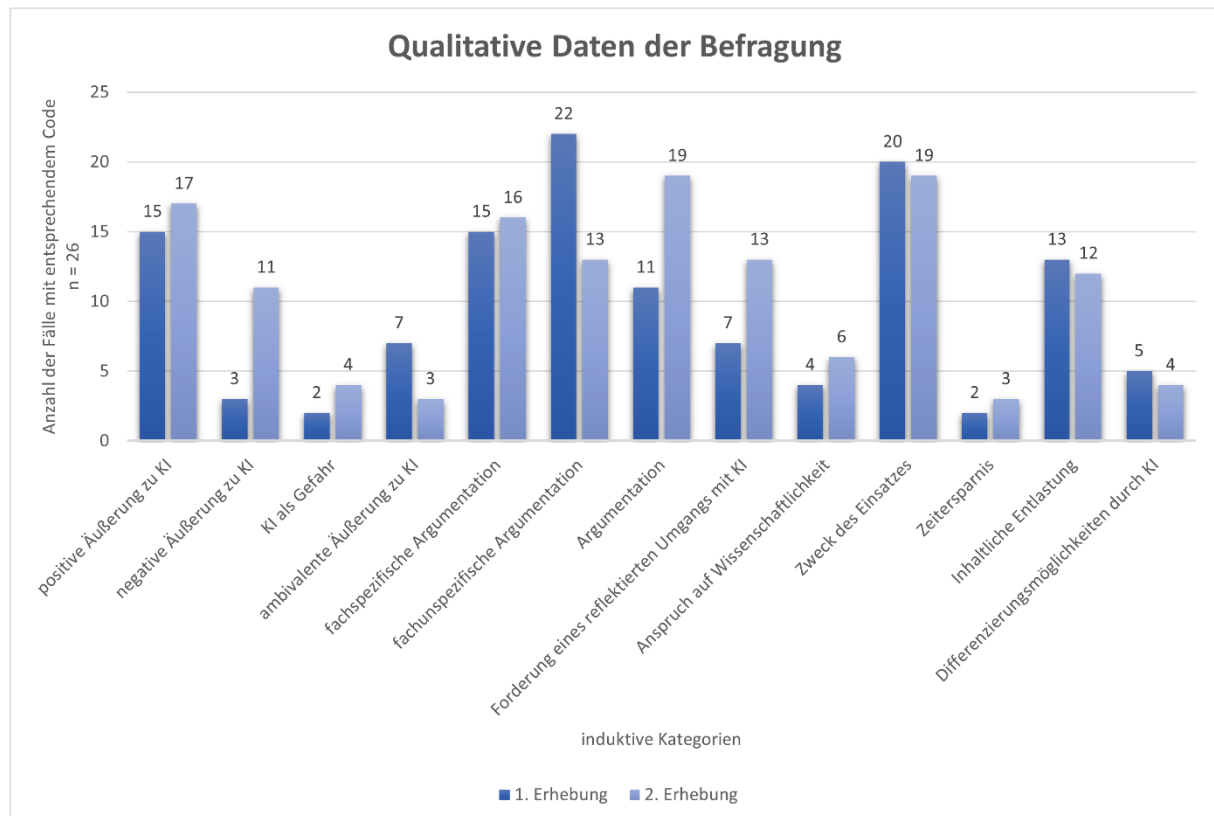


Abbildung 3: Qualitative Daten der Befragung

Eine grundlegend positive Einstellung zu KI zeigen in der ersten Befragung 65 % der Studierenden. Nur eine Person äußert sich negativ; 30 % der Studierenden geben an, eine ambivalente Haltung zu haben. Der Befund der zweiten Befragung unterscheidet sich maßgeblich davon: Zwar steigt die Anzahl der grundlegend positiv eingestellten Studierenden leicht, doch zeigen sich die zuvor ambivalent eingestellten Studierenden nun entschieden negativ eingestellt gegenüber dem Einsatz von KI. Lediglich eine Person weist eine ambivalente Haltung aus.

Hinsichtlich der Unterrichtsideen und der jeweiligen Argumentation in der offenen Schreibaufgabe konnten in beiden Befragungen fachspezifische sowie fachunspezifische Anwendungsstrategien ausgewiesen werden. An dieser Stelle wird die Qualität der Vorschläge, die die Studierenden machen, nicht bewertet, d. h. dass der Code „fachspezifische Anwendungsstrategie“ auch bei anscheinend nur oberflächlich mit geschichtsdidaktischen Überlegungen angereicherten Äußerungen vergeben wurde. In der ersten Befragung wiesen die Argumentationen von zwölf Studierenden fachspezifische Überlegungen aus. Nach dem Seminar argumentierten 18 der Teilnehmenden fachspezifisch. Fachunspezifische Überlegungen stellten dagegen in der ersten Befragung 22 Studierende an; in der zweiten Befragung waren es noch 13. Dabei waren Überschneidungen möglich, d. h. dass bei einer Person sowohl fachspezifische als auch fachunspezifische Überlegungen codiert wurden.

Zu Beginn des Semesters haben sieben Studierende einen reflektierten Umgang mit KI beim Einsatz im Geschichtsunterricht gefordert; zum Ende des Semesters postulierten dies doppelt so viele Teilnehmende. Auch ein Anspruch auf Wissenschaftlichkeit wurde in beiden Befragungen an die einzusetzende KI herangetragen (E1: 4/E2: 6). Dieses Ergebnis ist an die bereits erwähnte Studie von Garrel et al. (2023) zurückzubinden. Die Forschenden konnten zeigen, dass viele Studierende großes Interesse an dem „wissenschaftlichen Output“ hatten, das sie zu generieren glaubten. Gleichzeitig war für diese Studierenden weniger relevant, wie ein solches zustande kommt (Garrel et al., 2023). Dies ist auch in der vorliegenden empirischen Studie der Fall, wie im Verlauf der Analyse der qualitativen Daten gezeigt werden kann. Als weitere fachunspezifische Einsatzzwecke von KI wurden Zeitersparnis (E1: 2/E2: 3), inhaltliche Entlastung (E1: 13/E2: 12) und Differenzierungsmöglichkeiten durch KI (E1: 5/E2: 4) angegeben.

Als fachspezifische Anwendungsstrategien konnten Folgende identifiziert werden:

Strategie 1: Glaubwürdigkeitsprüfung

Diese Strategie ist codiert worden, wenn Studierende vorgeschlagen haben, im Geschichtsunterricht eine Quellenkritik durchführen zu lassen, wobei dies einerseits KI-gestützt, andererseits „klassisch“ erfolgen sollte, um die von der KI generierten Ergebnisse zu überprüfen. Zudem ist diese Strategie codiert worden, wenn Teilnehmende großen Wert auf die Prüfung der empirischen Triftigkeit legten und sie in ihren Vorschlag für KI-gestützten Geschichtsunterricht einbrachten. In beiden Anwendungsfällen liegt der Schwerpunkt auf der Prüfung der Glaubwürdigkeit. Diese Strategie ist von insgesamt 16 Teilnehmenden eingesetzt worden.

Strategie 2: Orientierung an Prinzipien historischen Denkens und Lernens

Wenn diese Strategie codiert wurde, haben sich die jeweiligen Lehramtsstudierenden in ihren Ausführungen explizit an Prinzipien historischen Denkens oder Lernens orientiert und ihre Überlegungen für den Einsatz von KI im Geschichtsunterricht danach ausgerichtet. Zentrale Konzepte, die berücksichtigt und genannt wurden, sind „Kompetenzen historischen Denkens“ (fünf Codierungen), „Historisches Lernen“ (zwei Codierungen), „Narrativität“ (zwei Codierungen), „Geschichtsbegriff“ (zwei Codierungen), „Perspektivität“ (eine Codierung), „Multiperspektivität“ (eine Codierung), „Handlungsorientierung“ (eine Codierung), „Sinnbildungen“ (eine Codierung) sowie Ideen, KI nutzbar zu machen, um den Konstruktcharakter von Geschichte deutlich zu machen (eine Codierung).

Strategie 3: Einsatz in Methodenfragen

Diese Strategie ist zweimal codiert worden. In beiden Fällen intendierten die Teilnehmenden, dass die Nutzung von KI ausschließlich im Sinne der Klärung von Methodenfragen geschehen solle. Dabei sollte beispielsweise erläutert werden, wie im Geschichtsunterricht eine Karikatur analysiert wird.

Auffällig ist, dass diese Strategien in der Systematik von Held (2024) hauptsächlich der Analyse von Materialien zuzuordnen sind. Selten schlagen Studierende vor, Bilder zu generieren. Eine Simulation von Interaktion kam gar nicht vor. Im Rahmen der ersten Strategie (Glaubwürdigkeitsprüfung) lassen sich die Überlegungen der Studierenden gut an die Ausführungen von Burkhardt & Neubert (2024) anschließen, die ausdrücklich darauf hinweisen, dass (u. a.) KI-generierte Erzählungen an Quellen zurückgebunden und Triftigkeiten geprüft werden müssten. Den Studierenden scheint diese Notwendigkeit zum Erhebungszeitpunkt bewusst gewesen zu sein.

Auf qualitativer Ebene geht aus den Daten hervor, dass der Einsatz von KI dabei in Teilen als naiv bezeichnet werden kann und bei der Umsetzung in vielen Fällen mutmaßlich zu problematischen Ergebnissen führen würde. Das Vorgehen wird als „in Teilen naiv“ charakterisiert, weil es sich in den meisten Fällen durch mangelnde Reflexion und Unbedarftheit auszeichnet. Grund dafür ist, dass viele Studierende nicht damit vertraut waren, wie generative KI-Tools funktionieren. Hier liegt die Ursache nicht bei den Studierenden selbst, sondern darin, dass der Erwerb von Kompeten-

zen im Umgang mit KI bisher schlicht nicht Teil ihres Studiums war. Unsere Interpretation der Daten wird gestützt durch die Angaben, die die Teilnehmenden auf die Frage nach dem Einsatz von KI im Rahmen des eigenen Studiums machten. So gaben Studierende etwa an, ChatGPT bei der Literaturrecherche oder beim Schreiben von Hausarbeiten einzusetzen. Eine solche Aufgabe kann ChatGPT jedoch aufgrund seiner Konzeptionsweise (noch) nicht sinnvoll bearbeiten. Dieser Befund ist in aller Vorsicht im Kontext der Ergebnisse von Lindner & Romeike (2019) zu betrachten. Sie hatten gezeigt, dass das Wissen vieler Lehrkräfte, die sie untersucht hatten, als defizitär zu bezeichnen sei (Lindner & Romeike, 2019).

Das Wissen der Lehramtsstudierenden konnte in dieser Studie nicht erhoben werden. Dennoch sind Selbsteinschätzungen abgefragt worden, die zeigen, dass in dieser Hinsicht eine große Heterogenität unter den Befragten herrscht.

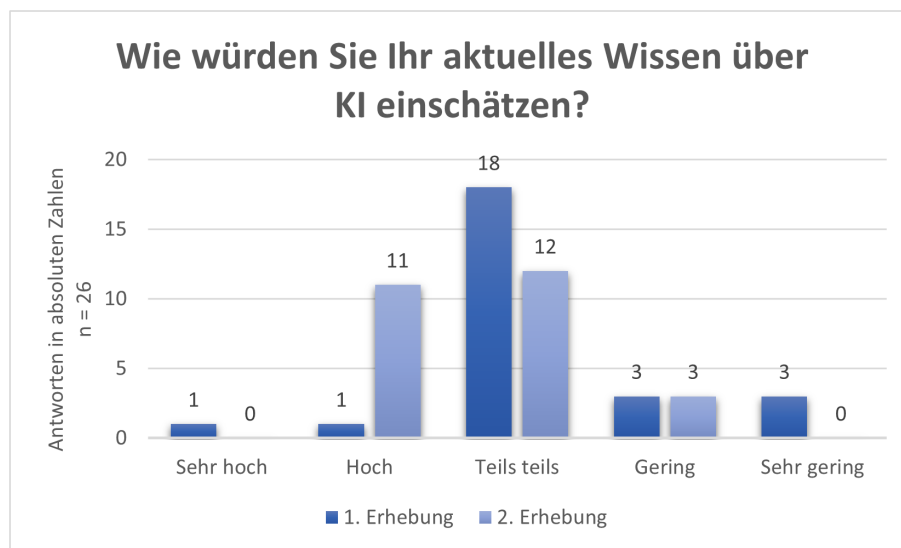


Abbildung 4: Ergebnisse aus der 1. und 2. Erhebung

Mit Blick auf die Seminarkonzeption ist es grundlegend als wünschenswert zu betrachten, dass circa die Hälfte der Teilnehmenden eine positive Veränderung hinsichtlich ihres Wissens über KI wahrzunehmen glauben.

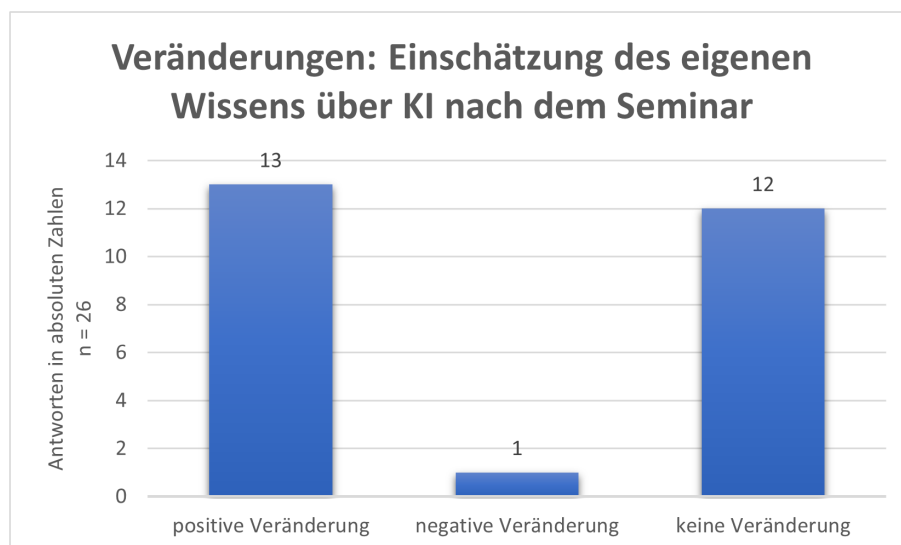


Abbildung 5: Ergebnisse aus der 2. Erhebung

6 Diskussion und Ausblick

Über das Semester hinweg ergaben sich Veränderungen in Bezug auf das Vorwissen über KI sowie die Anwendungsstrategien der Studierenden. Während zu Beginn des Semesters viele Studierende angaben, wenig Wissen über KI zu haben, fielen die Selbsteinschätzungen am Ende des Semesters deutlich positiver aus. Auch im Bereich der Anwendungsstrategien zeigt sich, dass die Studierenden sowohl ausführlichere Ideen für den Einsatz von KI im Geschichtsunterricht haben als auch stärker fachspezifisch und differenzierter argumentieren als noch zu Beginn des Semesters. Diese Veränderungen lassen sich kausal jedoch nicht ausschließlich auf die geschichtsdidaktischen Proseminare zurückführen, da anzunehmen ist, dass auch andere Lehrveranstaltungen zu diesen Veränderungen beigetragen haben. Darüber hinaus erwähnen die Studierenden selbst andere Lehrveranstaltungen, in denen der Einsatz von KI reflektiert wurde.

Dennoch wird anhand der beobachteten Veränderungen deutlich, wie viel ein einziges Semester hinsichtlich der Anwendungsstrategien zu KI bewirken kann. Auf Grundlage neuer Erfahrungen sowie der Reflexion im Seminar ergaben sich signifikante Veränderungen. So zeigen die Ergebnisse dieser Studie, dass die Studierenden bereit sind, sich zu KI und deren Anwendungsfeldern im schulischen Kontext zu verhalten. Die Reaktion erfolgt dabei vor dem Hintergrund des jeweiligen Vorwissens sowie der Erfahrungen mit KI. Das bedeutet, dass es Impulse in der (fachspezifischen) Hochschullehre braucht, um den Studierenden einen professionellen Umgang mit KI und den damit verbundenen Anforderungen und Herausforderungen zu ermöglichen.

Auf Ebene der Hochschullehre erscheint es daher grundsätzlich notwendig, einen reflektierten Einsatz von KI in die Lehre zu integrieren. Das setzt jedoch sowohl einen kompetenten Umgang mit KI vonseiten der Lehrenden als auch klar umrissene rechtliche Vorgaben für den Einsatz von KI in Lehre und Studium voraus. Gleichzeitig erfordert der Umgang mit KI Flexibilität und Innovation in der Lehre – ein Changieren zwischen fachwissenschaftlichen/-didaktischen Erkenntnissen einerseits und technischen Entwicklungen andererseits erscheint maßgeblich für den professionellen Einsatz von KI in der Hochschullehre. Diese Aufgabe wird, sicherlich aus verschiedenen Gründen, dennoch viel zu häufig den Lehrenden selbst überlassen, ohne Kapazitäten oder Angebote für die entsprechende Weiterbildung bereitzustellen.

Für den fachspezifischen Einsatz von KI erscheint es sinnvoll, diesen praxisorientiert und veranstaltungsübergreifend zu reflektieren. Hierbei sind sowohl fachwissenschaftliche als auch fachdidaktische Erkenntnisse zu beachten. Die Reflexion des Einsatzes von KI-Anwendungen sollte dabei zwingend theoriegeleitet erfolgen. Nur so kann ein fachspezifischer Einsatz von KI gewährleistet und die Anwendung von KI um ihrer selbst Willen vermieden werden. Der Einsatz von KI sollte dabei stets vom Ziel aus geplant werden. Die theoriegeleitete Reflexion von KI zieht gleichzeitig eine Diskussion und Weiterentwicklung von Theorie nach sich. Bezogen auf die Geschichtsdidaktik wirft die Anwendung von KI beispielsweise die Frage nach Perspektivität auf: Was zeichnet eigentlich eine Perspektive aus und hat KI selbst eine Perspektive? Neben verschiedenen theoretischen Überlegungen erscheint es sinnvoll, den Einsatz von KI im Kontext von Hochschullehre und Schulunterricht auch empirisch zu untersuchen. Bezogen auf den Geschichtsunterricht stellt sich u. a. die Frage, inwiefern historische Lernprozesse unter Berücksichtigung von KI stattfinden (können).

Anmerkungen

Wir danken allen Teilnehmenden des Seminar- und Forschungsprojekts. Allen voran möchten wir unseren Studierenden für das großartige und spannende Semester sowie die Teilnahme an der Studie danken. Ohne Ihr Engagement und Ihre Unterstützung wäre dieses Projekt nicht möglich gewesen! Der Besuch im MINT-Schülerlabor der Universität Paderborn im Rahmen der Praxisphase war zentraler Bestandteil des Seminarkonzepts und hat damit auch entscheidend zum Gelingen unseres Forschungsprojekts beigetragen. Ein besonderer Dank gilt Tobias Blomberg, der uns bei der Pla-

nung und Durchführung der Praxisphase unterstützt hat. Außerdem danken wir den Geschichtslehrern Sebastian Kelling und Henrik Brings (vom Gymnasium Schloß Neuhaus), die uns mit ihren Schüler:innen im Schülerlabor besucht haben. Ein großer Dank gilt außerdem Alexandra Krebs, die uns bei der Durchführung dieses Forschungsprojekts beratend zur Seite stand!

Literatur

- Abdelghani, R., Wang, Y.-H., Yuan, X., Wang, T., Sauzéon, H. & Oudeyer, P.-Y. (2022). GPT-3-driven pedagogical agents for training children's curious question-asking skills. https://www.researchgate.net/publication/365786306_GPT-3-driven_pedagogical_agents_for_training_children's_curious_question-asking_skills
- Aeppli, J., Gasser, L., Gutzwiller, E. & Tettenborn Schärer, A. (2014). *Empirisches wissenschaftliches Arbeiten. Ein Studienbuch für die Bildungswissenschaften*. Klinkhardt.
- Almendag, E. (2023). The effect of chatbots on learning: a meta-analysis of empirical research. *Journal of Research on Technology in Education*. <https://doi.org/10.1080/15391523.2023.2255698>
- Barricelli, M. (2011). Historisches Erzählen. Was es ist, soll und kann. In O. Hartung, I. Steininger & T. Fuch (Hrsg.), *Lernen und Erzählen interdisziplinär* (S. 61–82). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-93478-5_4
- Barricelli, M. (2016). Problemorientierung. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 78–90). Wochenschau Verlag.
- Bergmann, K. (2016). Lebenswelt- und Zukunftsbezug. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 91–112). Wochenschau Verlag.
- Bergmann, K. (2016). Multiperspektivität. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 65–77). Wochenschau Verlag. <https://doi.org/10.46499/861>
- Bhattacharya, A., Dutta, S., Dutta, P. & Samanta, D. (2023). *Innovations in Data Analytics*. Springer. <https://doi.org/10.1007/978-981-99-0550-8>
- Borries, B. (2013). Wissenschaftsorientierung. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 30–48). Wochenschau Verlag.
- Burkhardt, H. & Neubert, A. (2024). Historisches Lernen mit Künstlicher Intelligenz? Überlegungen und Anregungen zum Umgang mit generativen Sprachmodellen wie ChatGPT im Geschichtsunterricht. *Geschichte für heute*, 1, 71–86.
- Deng, X., Yu, Z. (2023). A Meta-Analysis and Systematic Review of the Effect of Chatbot Technology Use in Sustainable Education. *Sustainability*. <https://doi.org/10.3390/su15042940>
- Gerell, J., Mayer, J. & Mühlfeld, M. (2023). *Künstliche Intelligenz im Studium. Eine quantitative Befragung von Studierenden zur Nutzung von ChatGPT & Co.* Hochschule Darmstadt. https://doi.org/10.48444/h_docs-pub-395
- Hattie, J. (2008). *Visible Learning. A Synthesis of over 800 Meta-Analyses relating to Achievement*. Routledge.
- Hattie, J. (2024). *Visible Learning 2.0. Deutschsprachige Ausgabe von „Visible Learning: The Sequel“* besorgt von Stephan Wernke und Klaus Zierer. Schneider Verlag Hohengehren.
- Held, O. (2024). *ChatGPT im Geschichtsunterricht*. Wochenschau Geschichte. <https://doi.org/10.46499/2269>
- Hopf, C. & Schmidt, C. (1993). *Zum Verhältnis von innerfamiliären sozialen Erfahrungen, Persönlichkeitsentwicklung und politischen Orientierungen: Dokumentation und Erörterung des methodischen Vorgehens in einer Studie zu diesem Thema*. Institut für Sozialwissenschaften der Universität Hildesheim. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-456148>
- Krebs, A. (2024). *Geschichten im digitalen Raum: Historisches Lernen in der „App in die Geschichte“*. De Gruyter. <https://doi.org/10.1515/9783111349558>
- Kuckartz, U. & Rädiker, S. (2024). *Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und künstlicher Intelligenz*. Beltz.
- Lee, H. (2022). A meta-analysis of the effects of using AI chatbot in Korean EFL education. *Studies in English Language & Literature*. <http://www.aellk.or.kr/datax/thesis/111995043391172.pdf>
- Lindner, A. & Romeike, R. (2019). Teachers' Perspectives on Artificial Intelligence. *ISSEP*, 22–29.
- Mayer, T. (2023). Chatten mit historischen Persönlichkeiten. Kompetenzcheck mit ChatGPT. *Geschichte Lernen*, 213, 62–63.

- Meyer-Hamme, J. (2018). Was heißt „historisches Lernen“? Eine Begriffsbestimmung im Spannungsfeld gesellschaftlicher Anforderungen, subjektiver Bedeutungszuschreibungen und Kompetenzen historischen Denkens. In T. Sandkühler, C. Bühl-Gramer, A. John, A. Schwabe & M. Bernhardt (Hrsg.), *Geschichtsunterricht im 21. Jahrhundert* (S. 75–92). Vandenhoeck & Ruprecht. <https://doi.org/10.14220/9783737008914.75>
- Pandel, H.-J. (2013). Quelleninterpretation. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 152–171). Wochenschau Verlag.
- Präsidium der Universität Paderborn (31.05.2022). Besondere Bestimmungen der Prüfungsordnung für den Bachelorstudiengang Lehramt an Gymnasien und Gesamtschulen mit dem Unterrichtsfach Geschichte an der Universität Paderborn. Prüfungsordnung Geschichte, Bachelor o. E., vom 7.2.22. *Verkündungsblatt*. <https://plaz.uni-paderborn.de/lehramtsstudium/pruefungsverfahren/pruefungsordnungen#c854193>
- Rüsen, J. (1983). *Historische Vernunft. Grundzüge einer Historik 1: Die Grundlagen der Geschichtswissenschaft*. Vandenhoeck & Ruprecht.
- Rüsen, J. (2003). Was ist Geschichte? In Ders. (Hrsg.), *Kann gestern besser werden? Essays zum Bedenken der Geschichte* (S. 109–139). Kulturverlag Kadmos.
- Schmohl, T., Watanabe, A. & Schelling, K. (2023). Künstliche Intelligenz in der Hochschulbildung: Chancen und Grenzen des KI-gestützten Lernens und Lehrens. In T. Schmohl, A. Watanabe & K. Schelling (Hrsg.), *Künstliche Intelligenz in der Hochschulbildung* (S. 7–26). transcript. <https://doi.org/10.1515/9783839457696-001>
- Schreiber, W. (2008). Ein Kompetenz-Strukturmodell historischen Denkens. *Zeitschrift für Pädagogik*, 54 (2), 198–212.
- Schönemann, B. (2012). Geschichtsbewusstsein. Theorie. In M. Barricell & M. Lücke (Hrsg.), *Handbuch Praxis des Geschichtsunterrichts Bd. 1* (S. 98–111). Wochenschau Verlag.
- Ständige Wissenschaftliche Kommission (2024). *Large Language Models und ihre Potenziale im Bildungssystem. Impulspapier der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. SWK.
- Ulrich, I. & Heckmann, C. (2017). Taxonomien hochschuldidaktischer Designs und Methoden aus pädagogisch-psychologischer Sicht samt Musterbeispielen aus der aktuellen Forschung. *die hochschullehre*, 3, 1–28.
- Völkel, B. (2016). Handlungsorientierung. In U. Mayer, H.-J. Pandel & G. Schneider (Hrsg.), *Handbuch Methoden im Geschichtsunterricht* (S. 49–64). Wochenschau Verlag.
- Wannemacher, K. & Bodmann, L. (2021). *Künstliche Intelligenz an den Hochschulen. Potenziale und Herausforderungen in Forschung, Studium und Lehre sowie Curriculumentwicklung*. Hochschulforum Digitalisierung.
- Watanabe, A. (2023). Studierende im KI-Diskurs. Wie Studierende in einem Workshopformat über KI-gestütztes Lehren und Lernen angeregt werden. In T. Schmohl, A. Watanabe & K. Schelling (Hrsg.), *Künstliche Intelligenz in der Hochschulbildung* (S. 99–118). transcript. <https://doi.org/10.1515/9783839457696-006>
- Wegener, J. (2023). Logiken künstlicher Intelligenz verstehen. Herausforderungen und Möglichkeiten der Integration von ChatGPT in den Geschichtsunterricht. *Geschichte Lernen*, 214, 63–64.
- Wu, R. & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13334>
- Zhang, S., Shan, C., Lee, J., Che, S. & Kim, J. (2023). Effect of chatbot-assisted language learning: A meta-analysis. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-11805-6>
- Zierer, K. (2024). *ChatGPT als Heilsbringer? Über Möglichkeiten und Grenzen von KI im Bildungsbereich*. Waxmann.

Autorinnen

Isabel Elsner. Universität Paderborn, Historisches Institut, Paderborn, Deutschland;
E-Mail: ielsner@mail.uni-paderborn.de

Franziska Pilz. Universität Paderborn, Historisches Institut, Paderborn, Deutschland;
E-Mail: franziska.pilz@uni-paderborn.de



Zitiervorschlag: Elsner, I. & Pilz, F. (2026). Dem geschichtsspezifischen Einsatz von KI auf der Spur. Eine hochschuldidaktische Untersuchung der Anwendungsstrategien von Lehramtsstudierenden im Kontext von Geschichtsunterricht. *die hochschullehre*, Jahrgang 12/2026. DOI: 10.3278/HSL2604W. Online unter: wbv.de/die-hochschullehre