

Forschungsethik und Künstliche Intelligenz (KI)

Vom handlungspraktischen Zugriff zu sich ändernden
Wissenschaftspraktiken im Kontext von
Erziehungswissenschaft¹

Juliane Ahlborn, Sandra Hofhues, Denise Klinge & Tobias Röhl

Abstract

Research Ethics and Artificial Intelligence (AI)

From Practical Application to Evolving Scientific Practices in
Educational Science

This article focuses on how educational science engages with the adoption and application of AI. By analyzing the use of specific tools, the authors reflect on scientific practice, ultimately concluding that AI is now consistently consulted within educational research processes, which calls for a broader commitment to research ethics.

Verwicklungen von Forschungsethik und KI

Der vorliegende Beitrag betrachtet angesichts der jüngeren medientechnischen Entwicklungen die Verwicklungen von Forschungsethik und KI. Dazu rückt die Frage in den Mittelpunkt, wie sich Erziehungswissenschaft zur Aneignung und Anwendung von KI verhält. So bewegt sich auch das Fach zwischen der Reflexion technologischer Entwicklungen rund um KI und einer normativen Beurteilung jener Ambivalenzen, die Digitalisierung deutlich macht. Herausforderungen entstehen somit nicht bloß durch die Nutzung von KI-basierten Anwendungen und Software-Systemen, sondern auch durch ihre Aneignung in der Breite. Die Ethikkommission der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE) begründet dies damit, dass Bildungsinstitutionen ein „hochrisikobehafteter Bereich“ sind, „in denen der Einsatz von KI-Systemen begründet, bewertet und auch reguliert werden muss“ (o. J.).

Dabei erweist sich KI als übergeordneter Begriff für technische Anwendungen jeder Art, die als Software maschinell, automatisiert und algorithmisch

1 Für ihre Unterstützung im Zuge der Beitragserstellung danken wir der Ethik-Kommission der DGfE, insbesondere Stefan Aufenanger, deren Anliegen rund um Forschungsethik und KI den Anlass zur weiteren Befassung geboten haben, sowie Gabi Reinmann, die uns als Critical Friend unterstützt hat.

lern- und anpassungsfähig sind. Schon allein deswegen gilt es als Erziehungswissenschaftler*in stets näher zu beurteilen, ob KI-Anwendungen in Forschung und Lehre ethisch unbedenklich eingesetzt werden können. Werden datenschutzrechtliche Bestimmungen unter Einsatz von Computertechnologien eingehalten, wenn Anwendungen wie selbstverständlich bei der Datenerhebung und -auswertung eingesetzt werden? Wie ist der Einsatz unter Gesichtspunkten des EU AI Acts zu beurteilen? Interessant ist einerseits, dass sich Fragen guter wissenschaftlicher Praxis auf diese Weise noch mehr mit forschungsethischen Aspekten vermischen (dazu auch DFG 2022). Andererseits verändern sich die Wissenschaftspraktiken selbst, sobald der Einsatz von KI mit dem Alltagshandeln als Erziehungswissenschaftler*in verweben ist. Und das bleibt ambivalent, wenn konkrete Tools z. B. für Recherchen sowohl Vor- als auch Nachteile bieten. So sind die daraus hervorgehenden Recherche-Ergebnisse immer nur wahrscheinliche Resultate, die auf probabilistischen Modellen beruhen. Was bedeutet das für die Erziehungswissenschaft?

Forschungsethik und KI: Anrufungen im Kontext einer datafizierten Erziehungswissenschaft

ChatGPT von OpenAI, Copilot von Microsoft oder Gemini von Google unterstützen auch Erziehungswissenschaftler*innen darin, forschungs- und lehrbezogene Themen zu strukturieren. Thematische Recherchen, Gliederungen von Texten, auch Bilder oder Datenanalysen entstehen auf Basis von Prompts. Sie werden sukzessive der weiteren Nutzung zugeführt. Wissenschaftler*innen verwalten längst auch ihre Literatur oder ihre Daten/Quellen mithilfe von Software-Anwendungen oder erheben mit anderer Software Daten jeder Art, werten sie darin aus und/oder führen sie einer Sekundär- respektive Nachnutzung zu.² KI kommt somit im gesamten Forschungsprozess zum Einsatz (Aufenanger 2025, in diesem Band). Neben der Tatsache, dass sich eine KI gestützte Erhebung *und* Auswertung als gängige Praxis etablieren könnte, ist für erziehungswissenschaftliche Forschungen bedeutsam, ob und inwieweit das Programm cloudbasiert bzw. über das Internet arbeitet. Wie könnte hier eine Art „Übersetzung“ des abschließbaren Schrankes für die Forschungsunterlagen für den digitalen Raum aussehen?

Ein typisches Beispiel ist „f4x Spracherkennung“. F4 läuft zwar im Browser, und die Dateien werden dort hochgeladen. Sie werden aber nach Angaben des Anbietenden datenschutzkonform verwaltet. Dazu sind Serverstrukturen erforderlich, welche die Daten innerhalb von Deutschland ausweisen. Die Audiodatei

2 Daten sind im Konnex mit KI danach zu unterscheiden, ob sie gezielt und damit zum Zwecke weiterer Auswertung erhoben werden oder ob sie durch die Nutzung einzelner digitaler Anwendungen anfallen und dann in erziehungswissenschaftlicher Perspektivierung weiter ausgewertet werden.

wird unmittelbar nach der Transkription gelöscht, und die erstellten Textdokumente werden nur für einen kurzen Zeitraum zum Download bereitgestellt. Aus datenschutzrechtlicher Sicht wäre der Einsatz dieses Programms unproblematisch. Das Programm „MacWhisper (Whisper transcription)“ nutzt während der Transkription keine Internetverbindung und ist damit ebenso als unproblematisch zu betrachten. Hier ist jedoch zu beachten, dass die Audiodateien bzw. Transkripte nicht an Personen gelangen, die nicht autorisiert sind bzw. mit denen keine Vertraulichkeitsvereinbarung abgeschlossen wurde. Darüber hinaus gibt es eine Reihe von Anbietenden, die Transkriptionen rein online durchführen. Dazu gehört z. B. Microsoft Word, das die Audiodatei zur Transkription auf eigene Server hochlädt. Auch qualitative Analysesoftware wie etwa MAXQDA beinhaltet KI-gestützte Funktionen, die auf die entsprechenden Plattformen von OpenAI zugreifen und dementsprechend problematisch sind. Nicht zuletzt bieten Anwendungen wie ChatGPT die Transkription von Audiodateien an. In der Regel besteht bei solchen kostenlosen Programmen keine Möglichkeit, die Nutzung zu Trainingszwecken zu unterbinden. Bei den kostenpflichtigen Modellen hingegen gibt es häufig eine Einstellung, die das Speichern der eingegebenen oder bei der Auswertung gewonnenen Daten einschränkt oder unterlässt. Eine weitere Alternative bietet die Möglichkeit, KI-Modelle ganz oder teilweise lokal bzw. innerhalb Europas zu hosten. Bei der Analyse qualitativer Daten kann beispielsweise auf Programme zurückgegriffen werden, deren KI-Komponenten nicht oder nur in eingeschränkter Weise mit den US-Servern von OpenAI kommunizieren (QualCoder, DokuMet QDA/AI).

Stets zu achten ist somit auf die Datenschutzbestimmungen von Anbietenden, die Standorte der Server und darauf, wo Daten gesammelt (und wieder gelöscht) werden. Zugleich steht die Erziehungswissenschaft vor der Anforderung, ihre forschungsethischen Prämissen auch auf KI-Technologien zu übertragen. Sie findet sich in einer Lage wieder, in der sie sich fortwährend ins Verhältnis zu aktuellen technologischen, aber auch regulatorischen Entwicklungen setzen muss und die Anforderungen an wissenschaftliche Transparenz (noch) weiterwachsen.

Von der Nutzung einzelner Tools zur Reflexion über Wissenschaft(spraxis)

In der Erziehungswissenschaft liegt der Fokus der Forschungsethik auf der Erhebung von Daten und damit auf der Art und Weise, wie im Rahmen der Erhebung und der sich daran anschließenden Datenauswertung der besonderen Schutzbedürftigkeit von Menschen nachgekommen wird. Im Einklang mit den jeweils aktuellen Datenschutzbestimmungen (derzeit etwa DSGVO) sind Informationsschreiben für Beteiligte der Forschung anzufertigen; auch Einwilligungserklärungen müssen von Beforschten sowie – im Falle von Kindern und

Jugendlichen – von Betreuungspersonen eingeholt werden. Sowohl Informationsschreiben als auch Einwilligungserklärungen verfolgen den gemeinsamen Zweck, Menschen über Erhebungsformate zu informieren und ihnen das Recht zuzusprechen, dass sie ihre Zustimmung zu einer Unterstützung jederzeit und auch ohne Angabe von Gründen zurückziehen können. Reflektieren wir in diesem Zusammenhang über konkrete Anwendungen von KI, tauchen in der Verschränkung mit Forschungsethik jedoch weitere und durchaus tiefgreifende Anschlussfragen für die Erziehungswissenschaft auf.

Personen, die im Zuge eines tradierten Methodenrepertoires der Erziehungswissenschaft – etwa im Kontext von Befragungen oder Beobachtungen – genauer erforscht werden, lassen sich noch vergleichsweise gut ein- und abgrenzen. Somit besteht die Option, ihnen das Recht auf Anonymität zuzugestehen. Ein solches Zugeständnis ist unter dem Einsatz von KI unmittelbar erschwert. Forschende übernehmen dafür Verantwortung, genauer zu definieren, wie ihre Forschungsdaten in technischen Anwendungen verarbeitet werden und inwiefern diese Umgebungen klar umgrenzt sind. Nehmen wir das Beispiel generativer Sprachmodelle, die Transkriptionssoftware und teilweise auch Social Bots wie ChatGPT dauernd zugrunde liegen: Werden diese mit Interview-Transkripten z. B. zu Zwecken inhaltsanalytischer Zusammenfassung gespeist, ist es wahrscheinlich, dass auch für das menschliche Auge anonymisierte Personen unter Einsatz digitaler Technologien im Internet reidentifiziert werden können. Gerade weil die hinter den Social Bots liegenden mathematischen Modelle durch ein ebensolches Speisen mit Daten öffentlich trainiert und stets besser werden, gelingt dieses Zusammenspiel. Das Beispiel führt Chancen und Grenzen gängiger Praxis unmittelbar vor Augen: Forschungsethisch ist demnach wichtig, exakt auszuweisen, dass und welche Anwendungen bei der Datenerhebung und -auswertung genutzt werden und in welcher Form sie zum Einsatz kommen. Genau aufzuzeigen ist, wie ganz unterschiedliche Datenformen darin im Einzelnen technisch verarbeitet werden und inwiefern diese „personenbezogen“ sind.

Wer bereits länger mit digitalen Technologien in der beschriebenen Form arbeitet, wird gleichsam zu dem Schluss gelangen, dass der Anspruch an die systematische und in der Hinsicht vollumfängliche Darstellung des Vorgehens kaum mehr realistisch ist. Zwar können vorgeschaltete Informationsboxen bei schriftlichen Befragungen zur Information über die Erhebung dienen: Sie klären über das eigene Vorgehen auf, problematisieren aber auch die Grenzen. Das scheint auch für den Einsatz digitaler Technologien geboten. So verweist der Ethikkodex der DGfE (2016) darauf, dass das jeweilige Forschungsinteresse nicht über der Schutzwürdigkeit von Personen stehen darf. Es gilt darüber hinaus das sogenannte Transparenzgebot, was die Erhebung und Auswertung gleichermaßen betrifft. Was aber, wenn man als Forschende selbst gar nicht mehr diese Transparenz hat?

Die geforderte Transparenz der Datennutzung betrifft nicht nur den Datenschutz und die Informationspflicht gegenüber den Beforschten, sondern auch die epistemische Praxis der Forschenden. Wenn Forschende nicht oder nur begrenzt nachvollziehen können, wie KI-Anwendungen auf Literatur zurückgreifen und sie zusammenfassen oder wie sie qualitative Daten analysieren, dann ist fraglich, inwiefern hier dem Anspruch an intersubjektive Nachvollziehbarkeit Genüge getan wird. Ob eine Zusammenfassung einen Artikel adäquat wiedergibt oder eine Analyse methodischen Standards genügt, kann nur dann vollumfänglich beantwortet werden, wenn man die entsprechenden Schritte selbst durchführt – was den Gewinn an Effizienz der KI-Anwendung wiederum zunichtemacht.

Fazit

Es dürfte deutlich geworden sein, dass der Umgang mit KI in die selbstverständlichen Praktiken abgesunken ist und sich diese Praktiken an „allgemeiner“ Redlichkeit sowie an guter wissenschaftlicher Praxis orientieren, an die auch die Deutsche Forschungsgemeinschaft mit ihren Leitlinien erinnert (DFG 2022). So setzt schon das Speichern von Literatur voraus, dass diese mindestens in Form von (Meta-)Daten online zugänglich ist, sprich Literatur respektive Quellen überhaupt gefunden werden *können*. Das wiederum verweist auf die Position der Forschenden und umfasst letztlich ein Bekenntnis zu Open Access. Zudem dürfte sich das Verständnis von Daten nicht auf konkrete Erhebungssituationen beschränken, sondern sukzessive ihre Vor- und Nachgeschichte in die Datenauswertung und/oder ihre Dokumentation einschließen. In dieser Perspektivierung möchte man anschließend fragen: Wie können gute wissenschaftliche Praktiken hier künftig aussehen?

Praktische Konsequenzen finden sich auch im Bereich der Lehre. Laut eines aktuellen Artikels in der Süddeutschen Zeitung würde auch Plagiatsoftware dazu führen, dass „Studierende [...] aus Angst übermäßig viel“ (Gebler 2024) zitieren. Somit ist die Hochschulbildung von den skizzierten Fragen gleichermaßen betroffen und könnte Antworten in Konzepten forschungsorientierter Lehre finden. Damit wäre auch darüber zu reflektieren, inwieweit neue Wissenschaftspraktiken innerhalb der Erziehungswissenschaft Aspekte der KI-Interaktion als Forschungsgegenstand adressieren und einhergehende Verantwortlichkeiten für Erziehungswissenschaftler*innen allgemein zum Gegenstand zu machen. All diese offenen Fragen rücken Themen wie die disziplinäre Verantwortung, persönliche Umgangsweisen, aber auch fachwissenschaftliche oder hochschulische Strukturen, Transformationsprozesse etc. ins Blickfeld – zudem verdeutlichen sie, dass KI innerhalb eines Forschungsprozesses in der Erziehungswissenschaft inzwischen wohl *immer* konsultiert wird.

Unsere *zentrale* Frage, wie sich Erziehungswissenschaft zur Aneignung und Anwendung von KI verhält, dürfte demnach um zahlreiche Fragen erweitert

werden – und sie müssten weiterer Reflexion zugeführt werden: Welche Verwendungsweisen von KI lassen sich bereits als Wissenschaftspraktiken lesen? Welche forschungspraktischen Implikationen hat dies für das gesamte Fach? Wenn man bei den skizzierten Verfahren bisher eher an empirisch forschende Erziehungswissenschaft denkt, schließt sich womöglich auch die Frage danach an, welche Praktiken eher Verfahren der Datenerhebung, welche jene der Auswertung betreffen. Die Zusammenschau zeigt allerdings, dass es auch kurz greifen könnte, die Fragen losgelöst voneinander zu betrachten. Stattdessen verdeutlicht die Verknüpfung von Forschungsethik und KI die Schwierigkeit, eine Metaperspektive in der Erziehungswissenschaft einzunehmen, insbesondere vor dem Hintergrund eines unaufhaltsamen und kontinuierlichen technologischen Fortschritts. Dies gilt umso mehr, als KI tiefgreifende methodologische, epistemologische und datenschutzrechtliche Veränderungen mit sich bringt.

Juliane Ahlborn, M.A., ist wissenschaftliche Mitarbeiterin an der Fakultät für Humanwissenschaften an der Universität Kassel.

Sandra Hofhues, Prof. Dr., ist Professorin für Mediendidaktik an der FernUniversität in Hagen.

Denise Klinge, Prof. Dr., ist Professorin für Erziehungswissenschaft mit dem Schwerpunkt Medienbildung an der Goethe-Universität Frankfurt.

Tobias Röhl, Prof. Dr., ist Professor für Digitales Lernen und Lehren am Zentrum Bildung und Digitaler Wandel an der Pädagogischen Hochschule Zürich.

Literatur

Aufenanger, Stefan (2025): Künstliche Intelligenz und Forschungsethik – eine pragmatische Sicht. In: *Erziehungswissenschaft* 36, 70, S. 83-88.

DGfE (2016): Ethikkodex der DGfE. <https://www.dgfe.de/dgfe-wir-ueber-uns/ethik-rat-ethikkodex> [Zugriff: 23. Februar 2025].

Deutsche Forschungsgemeinschaft (DFG) (2022): Leitlinien zur Sicherung guter wissenschaftlicher Praxis. Kodex. https://zenodo.org/records/6472827/files/kodex_leitlinien_gwp_dfg.1.1.pdf?download=1 [Zugriff: 23. Februar 2025].

Ethikkommission der DGfE (o. J.). *Forschungsethik und KI*. Unveröffentlichtes Papier.

Gebler, Victoria (2024): Plagiatsoftware. „Studierende zitieren aus Angst übermäßig viel“. In: *Süddeutsche Zeitung* vom 27. Dezember 2024. <https://www.sueddeutsche.de/wissen/studierende-unrecht-text-matching-software-li.3168328> [Zugriff: 23. Februar 2025].

Open Access © 2025 Autor*innen. Dieses Werk ist bei der Verlag Barbara Budrich GmbH erschienen und steht unter der Creative Commons Lizenz Attribution 4.0 International (CC BY 4.0).