

Empfehlung zur Sicherung der guten wissenschaftlichen Praxis beim Umgang mit Künstlicher Intelligenz

Beschlossen durch das Präsidium
der Leibniz-Gemeinschaft
am 29. November 2024

Herausgeber

Die Präsidentin der Leibniz-Gemeinschaft
Martina Brockmeier
Chausseestraße 111
10115 Berlin
info@leibniz-gemeinschaft.de

Lizenzhinweis

Alle Texte dieser Publikation sind, ausgenommen Zitate und soweit nicht anders angegeben, unter der Lizenz Namensnennung 4.0 International (CC BY 4.0) veröffentlicht. Den vollständigen Lizenztext finden Sie unter:

<https://creativecommons.org/licenses/by/4.0/deed.de>



Redaktion

Johannes Bronisch (Geschäftsstelle Leibniz-Gemeinschaft), Sabine Brünger-Weilandt (Leibniz-Ombudsgremium), Luzia Goldmann (Geschäftsstelle Leibniz-Gemeinschaft), Hans-Georg Joost (Leibniz-Ombudsgremium, Sprecher), Andreas Radbruch (Leibniz-Ombudsgremium)

Ansprechperson in der Geschäftsstelle

Johannes Bronisch
Chausseestraße 111
10115 Berlin
bronisch@leibniz-gemeinschaft.de

Beschlussdatum

Beschlossen durch das Präsidium der Leibniz-Gemeinschaft am 29. November 2024

Webseite

<https://www.leibniz-gemeinschaft.de/gute-wissenschaftliche-praxis>

DOI

10.5281/zenodo.14420893

Zitierweise

Leibniz-Gemeinschaft. (2024). Empfehlung zur Sicherung der guten wissenschaftlichen Praxis beim Umgang mit Künstlicher Intelligenz. Zenodo.

<https://doi.org/10.5281/zenodo.14420893>

Diese Empfehlung ist in deutscher wie englischer Sprache [verfügbar](#).

This recommendation is [available](#) in German and English.

Systeme „Künstlicher Intelligenz“ (KI) entwickeln sich dynamisch und haben in unterschiedlichem Umfang Einzug in die wissenschaftliche Praxis gehalten. Sie beeinflussen wissenschaftliches Arbeiten voraussichtlich langfristig und tiefgreifend. Die Erprobung und Ausschöpfung der Potenziale ebenso wie die Steuerung der Risiken von KI-Systemen erfordern deshalb eine stete Begleitung und kritische Prüfung auch im Hinblick auf die Regeln der wissenschaftlichen Integrität. Hierbei gilt es, etablierte Standards auch in neuen Wissenschaftspraktiken zu wahren.¹

Mit KI-Systemen sind in der folgenden Empfehlung Systeme generativer KI gemeint. Dies schließt insbesondere unterschiedliche Nutzungsweisen von Large Language Model-basierten Anwendungen, vorrangig der Text- und Codegenerierung, sowie deren Nutzung als Recherche- und Sprachbearbeitungsinstrumente ein, aber auch KI-Anwendungen zur Bildgenerierung und Datenanalyse.

Auf Vorschlag des Leibniz-Ombudsgremiums und beschlossen durch das Präsidium der Leibniz-Gemeinschaft soll die hier vorliegende Empfehlung den Wissenschaftlerinnen und Wissenschaftlern in der Leibniz-Gemeinschaft als Orientierungshilfe bei der Nutzung von generativen KI-Systemen im Forschungsalltag dienen. Zudem möchte sie den Austausch über die konkrete Ausgestaltung und Weiterentwicklung der wissenschaftsadäquaten Nutzung von KI-Systemen in fachspezifischer Perspektive innerhalb der und zwischen den Einrichtungen der Leibniz-Gemeinschaft, aber auch darüber hinaus, anregen und mitgestalten.

Empfehlungen an die Wissenschaftlerinnen und Wissenschaftler in der Leibniz-Gemeinschaft

- 1. Verantwortung:** Die Regeln der wissenschaftlichen Integrität und der guten wissenschaftlichen Praxis gelten uneingeschränkt auch dann, wenn im Forschungsprozess sowie bei der Projektierung und Beantragung von Forschungsvorhaben KI-Systeme zur Anwendung gelangen. KI-Systeme können dabei lediglich unterstützende Instrumente sein. Ihre Resultate haben als solche keine wissenschaftliche Validität und sind kein Ersatz für authentische Erkenntnis- und Entscheidungsprozesse. Die Nutzung von KI-Systemen mindert deswegen nicht die

¹ Siehe u. a. Leibniz-Gemeinschaft (2021): Leibniz-Kodex gute wissenschaftliche Praxis. URL: www.leibniz-gemeinschaft.de/kodex-gute-wissenschaftliche-praxis (25.10.24); Deutsche Forschungsgemeinschaft (2023): Stellungnahme des Präsidiums der Deutschen Forschungsgemeinschaft (DFG) zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG. URL: <https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellungnahme-praesidium-ki-ai-data.pdf> (25.10.24); European Commission, Directorate-General for Research and Innovation (2024): Living guidelines on the responsible use of generative AI in research. URL: https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf (25.10.24).

Verantwortung von Wissenschaftlerinnen und Wissenschaftlern. Ihnen obliegt die Sicherstellung der Einhaltung der wissenschaftlichen Standards, wozu auch die Reflexion über einen methodologisch und inhaltlich sinnvollen Einsatz von KI-Systemen gehört. Der mit einer wissenschaftlichen Autorschaft an Forschungsarbeiten und -ergebnissen möglicherweise verbundene Reputationsgewinn kann nur auf dieser Verantwortung beruhen.

2. Autorschaft: KI-Systeme erfüllen nicht die Kriterien wissenschaftlicher Autorschaft. Diese beruht auf einer nicht übertragbaren Verantwortung für Korrektheit, Adäquanz und Nachvollziehbarkeit aller Angaben in Veröffentlichungen und Anträgen. Daraus folgt die Verpflichtung wissenschaftlicher Autorinnen und Autoren, Auswahl und Einsatz von KI-Systemen so vorzunehmen, dass ihnen und anderen eine möglichst weitgehende Überprüfung der durch die KI generierten Resultate möglich wird. Das Ausmaß der Verwendung von KI sollte in Bezug zum Grad der Nachvollziehbarkeit stehen. Darüber hinaus empfiehlt sich eine frühzeitige Berücksichtigung der Vorgaben des avisierten Publikationsorganes bzw. der adressierten Forschungsförderorganisation bezüglich der Nutzung von KI-Systemen. Bei gemeinsamen Veröffentlichungen sollte eine klare Abstimmung hinsichtlich Art und Umfang der Anwendung von KI-Systemen unter allen Autorinnen und Autoren erfolgen und dokumentiert werden.² Die Nutzung von KI-Systemen zur Generierung oder Manipulation von Realbildern verschiedener Art wie Fotografien, Mikroskopiebildern, Röntgen-, CT-, MRT- und nuklearmedizinischen Bildern, Western-Blots- und Zellkultur-Abbildungen, Fluoreszenzbildern sowie Spektral-, Teleskop-, Luft- und Satellitenaufnahmen u.ä. ist unzulässig.³

3. Nachweise und Zitationsregeln: Die Nutzung von KI-Anwendungen⁴ im Forschungsprozess unterliegt den etablierten Nachweisanforderungen der guten wissenschaftlichen Praxis. Sie muss immer offengelegt werden. Eine solche Offenlegung sollte in der Regel mindestens eine klare Nennung der mit Hilfe von KI erstellten oder bearbeiteten Inhalte unter Angabe der genutzten Software inklusive Eigentümer, Version sowie eine vollständige Dokumentation der verwendeten *Prompts* mit Datierung umfassen. Die Verwendung von *Prompt-engineering-tools* ist ebenso nachzuweisen. Beispiele hierfür finden sich im Anhang zu dieser

² Siehe auch Leibniz-Gemeinschaft (2023): Empfehlung zur Sicherung der guten wissenschaftlichen Praxis bei Koautorschaften. Präsidium der Leibniz-Gemeinschaft. URL: https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/%C3%9Cber_uns/Gute_wissenschaftliche_Praxis/Empfehlung_Koautorschaften.pdf (25.10.2024).

³ Diese Aussage zielt nicht auf Forschung, die in Gegenstand oder Vorgehensweise lege artis entsprechende KI-Systeme verwendet.

⁴ Die Ausführungen beziehen sich auf KI-Anwendungen im Sinne von generativen KI-Systemen.

Empfehlung. Bei großflächigem KI-Einsatz bspw. zur Verbesserung der Lesbarkeit eines Textes sowie bei der Nutzung von KI in der Phase der Ideen- und Hypothesenentwicklung wird eine entsprechende Erwähnung im Rahmen der Beschreibung der zum Einsatz gekommenen Methoden oder an anderer geeigneter Stelle empfohlen.

- 4. Begutachtungen:** Der Umgang mit unveröffentlichten Forschungsergebnissen und Forschungsdaten sowie mit Förder- und Projektanträgen Dritter ist mit hoher Verantwortung verbunden und bedarf besonderer Umsicht. Eine Eingabe in und Bearbeitung solcher Daten mit KI-Systemen – bspw. in Begutachtungs- und Peer-review-Prozessen – ist unzulässig, solange diesem Vorgehen datenschutz- und urheberrechtliche Bedenken entgegenstehen. Hinzu kommen im Rahmen von Gutachtertätigkeiten weitere Sorgfalts- und Vertraulichkeitsanforderungen, u.a. aufgrund der damit oftmals verbundenen Karriererelevanz für Dritte. Die Verantwortung für Bewertungen in Begutachtungsprozessen muss ungemindert bei den Gutachterinnen und Gutachtern verbleiben. Wie im Forschungsprozess können KI-Anwendungen hierbei nur unterstützende Funktionen einnehmen. Gutachterinnen und Gutachter sowie Herausgeberinnen und Herausgeber sollten zudem bei Unklarheiten über die Art und Intensität der Verwendung von KI-Systemen in den ihnen vorgelegten Anträgen oder Forschungsergebnissen auf Klärung oder Nachbesserung bestehen.
- 5. Beachtung von Rechtskontexten:** Die Dynamik der Rechtsprechung sowie die Entwicklung der Vorgaben und Leitlinien von Publikationsorganen und Forschungsförderorganisationen bezüglich der Nutzung von KI-Systemen erfordert aufmerksame Beobachtung. Grundsätzlich ist von einer Nutzung von KI-Systemen immer abzusehen, wenn dabei die Gefahr der unbefugten Weitergabe von persönlichen oder sensiblen Daten, von Geschäftsgeheimnissen oder von Informationen, die einer Geheimhaltungspflicht unterliegen, besteht oder eine Verletzung von Urheberrechten droht. Die konkret möglichen Anwendungsbereiche für – insbesondere proprietäre – KI-Systeme in der wissenschaftlichen Arbeit stehen in enger Abhängigkeit von den jeweiligen Nutzungsbedingungen und dem geltenden rechtlichen Rahmen.
- 6. Forschungsethik:** Die Einhaltung forschungsethischer Standards ist auch bei der Nutzung von KI-Systemen zu gewährleisten. Resultate aus KI-Systemen dürfen nicht an die Stelle von ethischen Reflexionsprozessen und verantwortungsvoller Entscheidungsfindung durch Wissenschaftlerinnen und Wissenschaftler treten. Diskriminierungs-, Verfälschungs- und Halluzinationseffekte und ihre Verstärkung, die sich u.a. aus den Trainingsdaten und der Funktionsweise der KI-

Anwendungen ergeben, erfordern eine besondere Aufmerksamkeit. Zudem sollte der insgesamt ressourcenintensive Einsatz von KI aus Gründen der Nachhaltigkeit auch im Forschungskontext stets mit Augenmaß erfolgen.

Empfehlungen an die Einrichtungen der Leibniz-Gemeinschaft

- 7. Einrichtungsinterne Reflexionsprozesse und Kompetenzaufbau:** Angesichts der Dynamik der Entwicklung und Verbreitung von KI-Systemen wird den Leibniz-Einrichtungen empfohlen, interne Reflexionsprozesse zur Nutzung von KI im Forschungsprozess und den Aufbau notwendiger Kompetenzen strukturiert zu fördern und institutsübergreifend Verantwortlichkeiten zu definieren. Hierbei sollten vor allem die forschungsethischen Implikationen des Einsatzes von KI, die möglichen Folgen für die epistemische Situation des Wissenschaftlers und Autors und – disziplin- und fachabhängig – der Stellenwert von Text thematisiert und vor dem Hintergrund der Institutsstrategie und -kultur diskutiert werden. Die Herstellung institutsübergreifender Transparenz über den Einsatz von KI-Systemen sollte dabei zu den Zielen gehören. Der Aufbau von Kompetenzen zur informierten und umsichtigen Nutzung von KI sollte für alle Wissenschaftlerinnen und Wissenschaftler der Leibniz-Einrichtungen in geeigneten Formaten und unter Hinzuziehung von interner oder externer Expertise gefördert werden. Einrichtungsinterne Verpflichtungen zum Einsatz von KI-Systemen sind abzulehnen, es sei denn, diese selbst sind Gegenstand oder *lege artis* Teil der Vorgehensweise von Forschung.
- 8. Aufgaben der Ombudspersonen:** Die Ombudspersonen der Leibniz-Einrichtungen werden gebeten, die hier formulierten Empfehlungen einrichtungsspezifisch zu reflektieren, ihre Umsetzung zu begleiten und zu ihrer Weiterentwicklung beizutragen. Eventuelle Verstöße gegen die Regeln der guten wissenschaftlichen Praxis im Zusammenhang mit der Verwendung von KI-Systemen sind entsprechend den geltenden Regelungen zu behandeln. Den Ombudspersonen wird empfohlen, sich an der allgemeinen Debatte um die Weiterentwicklung der Standards der guten wissenschaftlichen Praxis im Hinblick auf den Einsatz von KI-Systemen zu beteiligen.
- 9. Beteiligung am Fachdiskurs:** Den Mitgliedseinrichtungen der Leibniz-Gemeinschaft wird nachdrücklich empfohlen, sich an der jeweils fachspezifischen Diskussion über den Einsatz von KI zu beteiligen. Leibniz-Einrichtungen sollten ihre Expertise und ihre spezifischen Schwerpunkte in Forschung und Infrastruktur einbringen, um den jeweiligen fachlichen *lege artis*-Standard im Sinne einer bestmöglichen Sicherung wissenschaftlicher Integrität mitzuprägen.

10. Allgemeine Informationssicherheitshinweise: Leibniz-Einrichtungen, die in ihrem Forschungsfeld die Möglichkeit von regelmäßiger und den Anforderungen der wissenschaftlichen Integrität genügender Anwendung von KI-Systemen wahrnehmen, sollten die Etablierung einheitlicher, einrichtungswweiter Anwendungen in Abstimmung mit den zuständigen Beauftragten prüfen. Dabei wird empfohlen, dies mit einem forschungsbezogenen Qualitätssicherungsprozess zu begleiten.

Anlage:

Empfehlungen zur Zitation bei der Nutzung generativer Sprachmodelle

Für die gängigen Zitationsstile APA, Chicago und MLA existieren bereits Empfehlungen für die Zitierung von Ergebnissen sprachgenerierender KI in wissenschaftlichen Texten. Ausführliche Erläuterungen finden sich auf den entsprechenden Webseiten.⁵ Zentrale Angaben für die transparente Nutzung von KI sind in allen Zitationsstilen Angaben zum **Eigentümer**, zur **Versionsnummer** und **Nutzungsdaten** sowie Ausweisung der durch KI-Systeme bearbeiteten Passagen sowie der genutzten *Prompts* bzw. Chatverläufe. Folgend finden sich zu allen drei Stilen Beispiele für wörtliche Zitate sowie bei MLA ein Beispiel für ein Bildzitat zur Orientierung:

a. APA-Style: <https://apastyle.apa.org/blog/how-to-cite-chatgpt>

Ein Hinweis auf die Verwendung von KI-Systemen sollte je nach Textgenre im Methodenteil oder der Einleitung erfolgen. Prompts sollten in den Text aufgenommen werden. Bei umfangreicheren Chatverläufen wird die Dokumentation der In- und Outputs über die zitierten Passagen in einem Anhang empfohlen. Alternativ lassen sich längere Chatverläufe auch über <https://sharegpt.com/> oder <https://aiarchives.org/> verlinken und teilen.

Zitationsbeispiel:

When prompted with “Is the left brain right brain divide real or a metaphor?” the ChatGPT-generated text indicated that although the two brain hemispheres are somewhat specialized, “the notation that people can be characterized as ‘left-brained’ or ‘right-brained’ is considered to be an oversimplification and a popular myth” (OpenAI, 2023).

Referenzangabe im Literaturverzeichnis:

OpenAI. (2023). *ChatGPT* (Mar 14, version xy) [Large language model].
<https://chat.openai.com/chat>

⁵ McAdoo, Timothy, American Psychological Association (2023): How to cite ChatGPT. URL: <https://apastyle.apa.org/blog/how-to-cite-chatgpt> (25.10.24); Modern Language Association of America (2023): How do I cite generative AI in MLA style?, URL: <https://style.mla.org/citing-generative-ai/> (25.10.24); University of Chicago (2024): Citation, Documentation of Sources. URL: <https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html> (25.10.24).

b. Chicago-Style: <https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html>

Um KI-Systeme in Literaturverzeichnisse aufnehmen zu können, wird nach Chicago-Style eine Verlinkung der vollständigen Chatverläufe bspw. über <https://sharegpt.com/> oder <https://aiarchives.org/> gefordert.

Zitationsbeispiel:

The following recipe for pizza dough was generated by ChatGPT: „Ingredients: 3 cups all-purpose flour, 1 tsp salt, 1 tsp sugar, 2 1/4 tsp active dry yeast (or 1 packet), 1 cup warm water (about 110°F or 43°C), 2 tbsp olive oil [...]“

Referenzangabe im Literaturverzeichnis:

- A. Fuß- bzw. Endnote: 1. ChatGPT, response to “Explain how to make pizza dough from common household ingredients,” OpenAI, March 7, 2023, version xy.
- B. Im Text-Angabe: „xxx“ (ChatGPT, March 7, 2023, version xy).

c. MLA-Style: <https://style.mla.org/citing-generative-ai/>

Auch im MLA-Style wird die Nutzung individueller Verlinkungen von Chatverläufen zur Dokumentation des Beitrags von KI-Systemen gefordert.

Zitationsbeispiel:

When asked to describe the symbolism of the green light in *The Great Gatsby*, *ChatGPT* provided a summary about optimism, the unattainability of the American dream, greed, and covetousness. However, when further prompted to cite the source on which that summary was based, it noted that it lacked “the ability to conduct research or cite sources independently” but that it could “provide a list of scholarly sources related to the symbolism of the green light in *The Great Gatsby*” (“In 200 words”).

Referenzangabe im Literaturverzeichnis:

“In 200 words, describe the symbolism of the green light in *The Great Gatsby*” follow-up prompt to list sources. *ChatGPT*, version xy, OpenAI, 9 Mar. 2023, chat.openai.com/chat.



Referenzangaben zu KI-generierten Bildern:

Fig. 1. "Pointillist painting of a sheep in a sunny field of blue flowers" prompt, *DALL-E*, version 2, OpenAI, 8 Mar. 2023, labs.openai.com/.