

Recommendations for safeguarding good research practice when using artificial intelligence

**Adopted by the Executive Board
of the Leibniz Association
29 November 2024**

Publisher

President of the Leibniz Association
Martina Brockmeier
Chausseestrasse 111
10115 Berlin
info@leibniz-gemeinschaft.de

Attribution

With the exception of citations, and unless otherwise indicated, all texts in this publication are published under the Attribution 4.0 International (CC BY 4.0) licence.

The full text of the licence can be found at:

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



Editors

Johannes Bronisch (Leibniz Headquarters), Sabine Brünger-Weilandt (Leibniz Ombuds Committee), Luzia Goldmann (Leibniz Headquarters), Hans-Georg Joost (Leibniz Ombuds Committee, Spokesperson), Andreas Radbruch (Leibniz Ombuds Committee)

Contact person at Leibniz Headquarters

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

Resolution date

Adopted by the Executive Board of the Leibniz Association on 29 November 2024

Web page

<https://www.leibniz-gemeinschaft.de/en/about-us/leibniz-integrity/good-scientific-practice-and-ombuds-services>

DOI

10.5281/zenodo.14420893

Cite this document

Leibniz Association. (2024). Recommendations for safeguarding good research practice when using artificial intelligence. 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.

“Artificial intelligence” (AI) systems are developing rapidly and have found their way into research practice to varying degrees. They are likely to have long-term and far-reaching impacts on research activities. Exploring and exploiting the potential of AI systems and controlling their risks are therefore activities that call for constant monitoring and critical scrutiny, including with regard to the rules of research integrity. Even with new research practices, it is essential to adhere to established standards.¹

In the following recommendations, “AI systems” refers to generative AI systems. In particular, these include various ways of using applications based on large language models, primarily for text and code generation, and their use as research and language-processing tools, but also AI applications for image generation and data analysis.

Proposed by the Leibniz Association’s Ombuds Committee, and adopted by the Executive Board, these recommendations are intended to guide the researchers of the Leibniz Association when using generative AI systems in their day-to-day research activities. In addition, it is hoped that the recommendations will spark and help structure debate within and between the institutes of the Leibniz Association, but also more broadly, about the concrete, scientifically appropriate ways that AI systems can be used and developed further in a subject-specific context.

Recommendations for researchers in the Leibniz Association

- 1. Responsibility:** The rules of research integrity and good research practice apply without limitation even when AI systems are used in the research process or in planning research projects and writing proposals. AI systems can only be supporting tools in these situations. As such, their results have no scientific validity and are no substitute for authentic cognitive and decision-making processes. Consequently, the use of AI systems does not reduce the responsibility of researchers. It is the researcher’s duty to ensure compliance with scientific standards, and this includes reflecting on a rational use of AI systems in terms of

¹ See e.g. Leibniz Association (2021): Leibniz Code for Good Research Practice. URL: www.leibniz-gemeinschaft.de/code-good-research-practice (25 Oct. 24); German Research Foundation (DFG) (2023): Statement by the Executive Committee of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) on the Influence of Generative Models of Text and Image Creation on Science and the Humanities and on the DFG’s Funding Activities. URL: <https://www.dfg.de/resource/blob/289676/89c03e7a7a8a024093602995974832f9/230921-statement-executive-committee-ki-ai-data.pdf> (25 Oct. 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 Oct. 24).

methodology and content. Any reputational gain associated with authorship of research papers and research findings can only be based on this responsibility.

- 2. Authorship:** AI systems do not meet the criteria for academic authorship. This is because responsibility for the accuracy, adequacy and verifiability of all information contained in publications and proposals is non-transferable. It follows that academic authors have an obligation to select and deploy AI systems in ways that enable them and others to verify the AI-generated results as thoroughly as possible. The extent to which AI is used should be commensurate with the degree of verifiability. Moreover, researchers are advised to check the requirements of the intended publication or targeted research funding organisation regarding the use of AI systems at an early stage. In the case of joint publications, there should be a clear, documented agreement between all the authors on how and to what extent AI systems are to be used.² The use of AI systems to generate or manipulate real images of various kinds, such as photographs, microscopy images, X-rays, CT scans, MRT scans, nuclear medicine images, western blot and cell culture images, fluorescence images and spectral, telescopic, aerial and satellite images, etc. is inadmissible.³

- 3. Referencing and citation rules:** Use of AI applications⁴ in the research process is subject to the established good research practice requirements regarding evidence. It must always be disclosed. The disclosure should generally contain at least a clear mention of the content created or processed with the help of AI, the software used, including the owner, the version, and full documentation of the prompts used with dates. The use of prompt engineering tools must also be referenced. Examples can be found in the Appendix to these recommendations. In cases where AI has been used on a large scale, for instance to improve the readability of a text, and where AI has been used in the ideas- and hypothesis-development phase, it is recommended that researchers mention this when describing the methods used or in another suitable section.

² See also Leibniz Association (2023): Recommendations for safeguarding good scientific practice in the context of co-authorship. Executive Board of the Leibniz Association. URL: https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/%C3%9Cber_uns/Gute_wissenschaftliche_Praxis/Recommendations_Co-Authorship.pdf (25 Oct. 2024).

³ This statement does not concern research that uses AI systems *lege artis* as its subject-matter or in its methodology.

⁴ Guidance relates to AI applications in the sense of generative AI systems.

- 4. Reviews and evaluations:** Handling unpublished research findings and research data and dealing with third-party funding applications and project proposals comes with a high level of responsibility and particular care must be taken in these situations. Entering and processing such data in AI systems – for example in the course of evaluation and peer-review processes – is inadmissible insofar as there are data protection and copyright objections to this approach. Moreover, reviewer activities entail additional care and confidentiality requirements, in part because they are frequently of relevance to the careers of third parties. Full responsibility for evaluations in review processes must remain with the reviewers. As in the research process, AI applications can only take on supporting functions. If there is doubt about the type and intensity of AI use in the proposals or research results submitted to them, reviewers and publishers should insist on clarification or rectification.

- 5. Compliance with legal contexts:** Changes to legislation and the development of rules and guidelines by publishers and research funding organisations regarding the use of AI systems need to be closely monitored. As a general rule, researchers should avoid using AI systems whenever there is a risk of unauthorised sharing of personal or sensitive data, trade secrets or confidential information, or of breach of copyright. The specific areas in which it is possible to use AI systems – especially proprietary AI systems – in research activities are heavily dependent on the conditions of use and the legal context in question.

- 6. Research ethics:** Compliance with ethical standards for research must be ensured, even when using AI systems. Results from AI systems must not be used in place of ethical reflection processes and responsible decision-making by researchers. Particular attention must be paid to discrimination, falsification and hallucination effects and the magnification of these effects, caused in part by the training data and functionality of AI applications. In addition, for reasons of sustainability, the use of AI, which is generally resource intensive, should always be kept in proportion, even in the research context.

Recommendations for the institutes of the Leibniz Association

- 7. Institutional reflection processes and developing expertise:** In view of the speed at which AI systems are developing and spreading, it is recommended that Leibniz institutes provide structured support for internal reflection processes on the use of AI in research and for the development of necessary expertise, and that they define responsibilities at institute level. In particular, this process should

address the implications of AI use for research ethics, the possible consequences for the epistemic situation of the researcher and author and – depending on the discipline and subject area – the value of text, and discuss these aspects in light of the institute’s strategy and culture. Creating institute-wide transparency regarding the use of AI systems should be one of the aims. Leibniz Institutes should support the development of expertise in informed, prudent AI use for all their researchers, in suitable formats and with the involvement of internal or external experts. Any institute-imposed requirement to use an AI system is to be rejected, unless the AI system is itself the subject of the research, or is *lege artis* a part of the research methodology.

- 8. Duties of the Ombudspersons:** The Ombudspersons of the Leibniz institutes are requested to reflect on the recommendations formulated here in terms of their own institution, to support their implementation and contribute to their further development. Any breaches of the rules of good research practice connected with the use of AI systems must be dealt with in line with the applicable regulations. It is recommended that the Ombudspersons take part in the general debate about the further development of standards of good research practice regarding the use of AI systems.

- 9. Participation in the professional discourse:** It is expressly recommended that member institutes of the Leibniz Association take part in the discussion about the use of AI in their respective fields. Leibniz institutes should contribute their expertise and their specialisms in research and infrastructure to the debate to help shape the relevant *lege artis* standards in the interests of ensuring optimum safeguards for research integrity.

- 10. General guidance on information security:** Leibniz institutes that seize the opportunity to make regular use of AI systems in their research field in ways that meet the requirements for research integrity should consult the responsible officers to investigate installing uniform, institute-wide applications. Institutes are advised to accompany this process with research-related quality assurance.

Appendix:

Recommendations for citing generative language models

The widely used APA, Chicago and MLA citation styles already contain recommendations for citing results from language-generating AI in academic texts. Detailed explanations can be found on the relevant websites.⁵ In all citation styles, key pieces of information to be provided for the transparent use of AI are the **owner** of the tool, the **version number** and the **dates used**, as well as an indication of the passages processed by AI systems and the prompts or chats used. Below are examples of citations for language-generating AI in all three styles for guidance, as well as an example of an image citation in MLA Style:

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

The use of AI systems should be mentioned in the method section or introduction, depending on the type of text. Prompts should be included in the text. In the case of longer chats, the recommendation is to put the AI inputs and outputs relating to the cited passages in an appendix. Alternatively, it is possible to create links and share longer chats using <https://sharegpt.com/> or <https://aiarchives.org/>.

Citation example:

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).

Reference:

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

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

⁵ McAdoo, Timothy, American Psychological Association (2023): How to cite ChatGPT. URL: <https://apastyle.apa.org/blog/how-to-cite-chatgpt> (25 Oct. 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 Oct. 24); University of Chicago (2024): Citation, Documentation of Sources. URL: <https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html> (25 Oct. 24).

To include AI systems in a bibliography or reference list, Chicago Style requires that a link be provided to the full chat, for example using <https://sharegpt.com/> or <https://aiarchives.org/>.

Citation example:

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 [...]”

Reference:

- A. Footnote or endnote: 1. ChatGPT, response to “Explain how to make pizza dough from common household ingredients,” OpenAI, March 7, 2023, version xy.
- B. In-text citation: “xxx” (ChatGPT, March 7, 2023, version xy).

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

MLA Style also requires the use of individual chat links to document the contribution of AI systems.

Citation example:

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”).

Reference:

“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.



Citing AI-generated images:

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/.