Discussion Hubs to Foster Research Integrity
www.ofdw.de/discussion-hubs
katrin.frisch@ofdw.de

FAQ ARTIFICIAL INTELLIGENCE AND RESEARCH INTEGRITY

This FAQ compiles a set of questions that we frequently receive in connection to artificial intelligence (AI) and research integrity (RI). The aim is to offer orientation in a fast-paced debate without being prescriptive. This FAQ does not constitute an official statement on the use of AI by the German Research Ombudsman (OfdW). Instead, it describes the status quo, contextualises existing recommendations, identifies gaps and provides further literature. The intended audience are researchers. The FAQ does not cover questions concerning student use of AI, since this is usually regulated by university policies, examination regulations (Prüfungsordnungen), declarations of originality (Selbstständigkeits-erklärungen) as well as individual decisions made by lecturers.

Compiled and written by Katrin Frisch

Last update: 6 November 2024

Contents

1. What is the general consensus on AI and RI?

Since the release of ChatGPT, policies and recommendations have reached a consensus on two aspects:

1) AI does not qualify for authorship as AI cannot take on responsibility for the contents of a manuscript nor can it agree to the final draft of a publication. Both of these aspects are common criteria for authorship in RI regulations.
2) The use of AI needs to be appropriately and transparently declared in the manuscript.

The specifics of declaring the use of AI differ in policies and practice or are still in need to be defined.

The following overview offers a summary of AI policies of the bigger publishers and publishing associations (last update: 6 November 2024)

	AI and authorship	Declaring the use of AI	AI-generated images	AI and Peer Review	Link AI Policy
Elsevier	AI ≠ author	Use of AI needs to be declared in manuscript	highly restricted	not allowed	https://www.elsevier.com/de-de/about/policies-and-standards/generative-ai-policies-for-journals
ICMJE	AI ≠ author	Use of AI needs to be declared in manuscript	not specified	highly restricted	https://www.icmje.org/icmje-recommendations.pdf
Springer Nature	AI ≠ author	Use of AI needs to be declared in manuscript	highly restricted	highly restricted	https://www.springernature.com/gp/policies/editorial-policies
Science	AI ≠ author	Use of AI needs to be declared in manuscript	highly restricted	not allowed	https://www.science.org/content/page/science-journals-editorial-policies#image-and-text-integrity
Taylor & Francis	AI ≠ author	Use of AI needs to be declared in manuscript	not allowed	highly restricted	https://taylorandfrancis.com/our-policies/ai-policy/?_ga=2.202881233.1550617214.1729762987-990921615.1729762987
Wiley	AI ≠ author	Use of AI needs to be declared in manuscript	not specified	highly restricted	https://authorservices.wiley.com/ethics-guidelines/index.html#22
WAME	AI ≠ author	Use of AI needs to be declared in manuscript	not specified	highly restricted	https://wame.org/page2.php?id=106

2. Why do I need to declare the use of AI?

Declaring the use of AI is in line with RI transparency standards (see guidelines 12 and 13 of the DFG Code of Conduct 'Guidelines for Safeguarding Good Research Practice'). Declaring the use of AI allows readers and reviewers to comprehend results, methods and the work process. Due to the sheer number of AI applications and their functions recommendations concerning the specifics of declaration may differ. Moreover, some aspects are still subject to discussion within the research community, for example the documentation of prompts (see question 6).

3. What does it mean to 'appropriately and transparently declare the use of AI'?

There is no single answer to this question. Most of the editorial policies only offer minimal information or no further specification. Some smaller publishing houses, e.g. Berlin Universities Publishing, have developed some more detailed recommendations. Existing recommendations differ regarding the extent and complexity of the disclosure, yet they all share a common core. This includes:

- name of AI application, including version, date of use, URL
- what the AI application was used for and how it was used

Further suggestions include specifying the member of the team that made use of AI (see [Hosseini et al. 2023](#)) and also detailed reflections on the technical functionalities and limitation of the AI used ([Resnik and Hosseini 2024](#)). Especially the latter exceeds common information required for declaring the use of AI, but it may be useful to reflect on the suitability of an AI application as well as its inherent weaknesses and limitations. In general, the declaration of the use of AI should do justice both to the needs of readers and reviewers as well as to the reality of working with AI. Both may be highly dependent on the discipline or field and thus should be subject to discussion within each research community. This should cover specifics on which types of AI use and which AI applications need to be declared ([see questions 7-11](#)).

4. In which part of the manuscript should the use of AI be declared?

This is not always specified in policies and recommendations. If specified, suggestions favour the methods section or the acknowledgments. Which part of the manuscript is the most suitable may also depend on field specific standards as well as the length/complexity of the declaration. Detailed declarations including the prompts used and the chat history may be added as a supplement (as suggested by [ASC Nano](#)).

5. Are there recommended citation styles for the declaration of AI?

If your institution or preferred journal does not have a recommended citation style, you may use existing style guides. The most established are the [American Psychological Association](#) (APA), [Chicago Style Manual](#) as well as the [Modern Language Association](#) (MLA). Please make sure that the citation style you use includes all the mandatory information requested by your institution/journal.

6. Do prompts need to be disclosed?

At present there is no consensus on this question. Of the big publishers only Science specifies that prompts need to be provided in the methods section. The [APA](#) and [MLA](#) style guides as well as the [Chicago Style Manual](#) do likewise recommend that prompts be provided, offering different suggestions on where and how they should be provided. Among researchers this is a more contentious subject: some question the usefulness of listing prompts as it neither corresponds to the actual (often iterative) use of AI nor offers increased transparency as AI-generated answers are not reproducible even with the exact same prompt. Whether prompts need to be disclosed should be discussed within research communities. Discussions should take into account which function the disclosure of prompts fulfils in the context of an individual publication but also within a certain field. Even if AI-generated results may not be reproducible, the prompts used by authors may offer readers insight into the work process.

7. Do I need to disclose the use of AI-generated code?

Editorial policies often do not specify how the use of AI in relation to code needs to be documented. The [World Association of Medical Editors](#) recommends that '[w]hen an AI tool such as a chatbot is used to [...] write computer codes, this should be stated in the body of the paper, in both the Abstract and the Methods section'. For transparency's sake, any use of AI in the writing or modifying of code should be documented in the manuscript. This is especially important if the code is made available to others.

8. Do I need to disclose the use of AI if I only use it to improve language or style?

Some editorial policies differentiate between different AI applications or their functions (for example generative AI on the one hand and tools that check grammar and spelling, like Grammarly, on the other). AI policies often do not cover the latter, i.e. their use does not need to be documented (see the AI policies by [Elsevier](#) or [Wiley](#)). The [DFG, likewise, specifies](#) that 'AI used that does not affect the scientific content of the application (e.g. grammar, style, spelling check, translation programmes) does not have to be documented.'

Scholars from different disciplines might disagree on this issue, depending on the role text and individual language use plays in publications. Especially in the humanities, style can be closely connected to individuals or schools of thought. In these fields, documenting the use of tools to improve style could be considered.

9. Do I need to disclose the use of AI if I use it for translations?

AI-generated translations are often not specifically mentioned in editorial policies. One exception is the DFG guideline on AI, which lists translation tools among those tools that do not require documentation (see also question 8). Likewise the BUP Guideline for Dealing with AI classifies translations tools as aids, that do 'not necessarily require explicit mention'. Authors should keep in mind that important information can get lost or distorted in AI-generated translations. Thus authors need to carefully check and proofread translated texts as they are responsible for any potential errors.

Moreover, translation can be considered an important personal contribution, skill or part of the self-conception in certain fields (e.g. Modern Languages). Translations of literary texts are especially seen as significant: 'a translation is the product of an individual handling of an original text. This needs to be undertaken responsibly, not only in the name of the translator but also in the name of the author of the original' (Original in German; VdÜ / A*ds / IGÜ – Offener Brief zur KI-Verordnung). In disciplines like Modern Languages or when translating literary texts it is recommended to document the use of translation tools.

10. Do I need to disclose the use of AI for inspiration?

Some recommendations exist on that matter. For example in the BUP Guidelines for Dealing with AI, which suggests that authors should in this case add a general note or a disclosure in the methods section. Authors should check whether the journal/publisher they are planning to submit to has any specifications on this matter. If not, authors should base their decision to disclose the use on discipline specific reading expectations as well as the role/extent of the AI-generated ideas for the publication.

11. Which tools are covered by the term AI?

In the general debate ChatGPT has become a synonym for generative AI, especially Large Language Models (LLM). Yet there are many AI applications that can be used in research (see, for example, the list of AI resources compiled by VK:KIWA or the Ithaka Product Tracker). With the ongoing technical development AI features are also added to existing applications, such as search engines. A clear demarcation between AI applications and non-AI applications might become increasingly difficult. It is possible to use a very broad definition of AI, as the EU AI Act does. Guidelines for researcher are mostly focussing on generative AI, especially LLM. AI Applications that are only used to improve language use and style fall into a grey area (see question 8). AI applications that can assist in the first steps of a research process (literature search, finding of hypotheses, literature review) are seldom specifically mentioned in existing guidelines. In the *Living Guidelines on the Responsible Use of Generative AI in Research* by the European Commission these AI applications are classified as "substantial use" and thus their use needs to be declared. Some editorial policies delineate which types of applications their guidelines cover. For example, Elsevier's AI Policy distinguishes between AI applications used during the scientific writing process, those for the research process, and other tools, such as spell checkers and reference managers: the policy includes different regulations for each. Authors should familiarise themselves in advance with the regulations that apply to them and, in case of doubt, be as transparent as possible when disclosing the use of AI.

12. Can I use AI for writing grant proposals?

Please refer to the information provided by the respective research funding organisation. Die Deutsche Forschungsgemeinschaft (DFG) permits the use of AI in grant proposals as long as it is appropriately disclosed (see their guidelines 'Use of Generative Models for Text and Image Creation in the DFG's Funding Activities' published in 2023). Moreover, the guidelines specify that '[i]n decision-making processes, the use of generative models in/for proposals submitted to the DFG is currently assessed to be neither positive nor negative' (DFG 2023). However, the use of AI is forbidden in the preparation of reviews (see also question 14).

13. Can I use AI to generate images?

Concerning AI-generated images, journals often have very restrictive policies (see also the overview of editorial policies above). The use of AI for the generation of images is usually limited to publications that specifically deal with the topic of AI. In these cases, similar to text generation, the use of AI to generate images has to be clearly marked. The publishing house Frontiers is one of the few exceptions as of now, that explicitly allows the use of AI-generated images, provided that authors disclose its use (see Frontiers Artificial intelligence: fair use and disclosure policy).

To my knowledge, no guidelines exist as of yet for the use of AI-generated images in other research output, such as presentations and posters. Researchers should discuss the issue with their project group or peers. AI-generated images that only serve illustrative purposes for use on presentation slides should be in line with research integrity. For graphics that visualise research results (like diagrams or flow charts) the use of AI could be permissible provided it is clearly marked. As with AI-generated text it is the responsibility of the researchers involved to check the results for accuracy.

14. Can I use AI when I review manuscripts or grant proposals?

In existing editorial policies, the use of AI in peer review is either subject to severe restrictions or not permitted at all (see overview in question 1). For reasons of confidentiality and data privacy, uploading a submitted manuscript (or grant proposal) into a generative AI application is generally not allowed. It should be noted that reviewing manuscripts/proposals 'has a prominent function in defining and shaping epistemic communities, as well as negotiating normative frameworks within such communities' (Hosseini und Horbach 2023). These key tasks should not be outsourced to an AI tool.

If editorial policies allow for a limited use of AI in peer review, it only applies to language post-processing (i.e. improving readability). Reviewers should check which requirements apply to them.

15. What should be taken into consideration when using AI in authorship teams?

Due to the relative novelty of many AI applications, it should be determined in authorship teams at the start of each project whether all authors agree to the use of AI tools and the extent of their use. This is especially relevant in trans- and interdisciplinary teams, since there is a higher chance of conflicting views on aspects such as text production (see also question 8). Teams can benefit from good internal documentation on the use of AI. In case of conflicts or breaches of research integrity it allows others to trace back the work process and genesis of the research manuscript. Authors publishing in teams may also consider following the suggestions by Hosseini et al 2023 to additionally document the member of the team who made use of AI (see also question 3).

16. Can I accidentally plagiarise other texts by using AI?

Since generative AI works in a stochastic way based on a huge set of training data, inadvertent plagiarism is not one of the common risks of AI. There are examples of well-known AI applications that can reproduce certain texts (almost) verbatim, despite their stochastic mode of operation (see Henderson et al. 2023 or the case of the New York Times versus OpenAI, e.g. Pope 2024). Yet in the studies, specific

prompting was used to achieve exactly these results. In some fields within the humanities, in which text and language use play a key role, some words or phrases may be attributable to individual famous theorists. If these terms are used without proper referencing this could be considered as missing citations or even as plagiarism (see Seadle: 'For the humanities, words matter. […] A stolen word is a stolen thought' (42)). Researchers are usually aware of the common discourses in their own fields. Thus especially researchers in transdisciplinary projects should carefully verify AI-generated output. In general, AI-generated text should only be used after extensive editing.

17. Which weaknesses and risks of AI should researchers keep in mind?

For generative AI especially hallucinating is a known weakness. Moreover, there are a number of further risks of known AI models. Those include missing or incorrect references, errors in direct quotes, fabricated quotes or references, outdated information as well as reproduction of bias and prejudice. The great percentage of English language sources (often of US-American provenance) should caution researchers using AI in different languages. A typology of risks and weaknesses of AI can be found in Oertner 2024 (in German). Authors bear responsibility for all potential errors and breaches produced by AI. Good prompting and a thorough review of the results can minimise the risks of use.

18. Can the use of AI be detected? How can I determine if a text is AI-generated?

There are a number of studies on AI detection tools, which have come to slightly different conclusions concerning the potential of these tools (Weber-Wulff et al. 2023, Gao et al. 2023). However, a review of the literature shows that generally there are no sufficiently reliable tools to detect AI-generated texts. Also human reviewers fail to reliably distinguish AI-generated texts from human-written ones. It is therefore not possible, at the moment, to definitively determine the use of AI in texts (unless a text contains certain recognisable phrases, see question 19).

19. I am a reviewer/editor and I suspect a text or parts thereof has/have been AI-generated, but the authors have not disclosed it. What am I supposed to do?

It depends what triggered your suspicion. There are some tell-tale phrases that an AI tool was used ('Certainly, here is an introduction for you', 'As an AI language model, I cannot…', 'as of my last knowledge update') as well as nonsense words or distorted font in AI-generated images. This would constitute sufficient proof that AI was used but not disclosed and thus a breach of existing editorial policies/recommendations. If this results in a desk rejection of the manuscript or the manuscript's authors be given the chance to rework the text and add the missing documentation should be discussed within each editorial team. Some editorial policies have set out a basic procedure what to do in these cases. If there are less concrete signs of AI use, such as style or certain words that are considered potential indicators (e.g. 'delve', 'meticulous', 'commendable'), the manuscript authors should be contacted to discuss the matter. Editors and reviewers need to keep in mind that neither the presence of the aforementioned potential indicator words nor any detection tool can reliably determine the use of AI. With regard to the DFG Code of Conduct, editors and reviewers should avoid any unfounded accusations that authors committed a breach against research integrity. Guideline 18 specifies that '[k]nowingly false or malicious allegations may themselves constitute misconduct'.

20. What happens when I am accused of having used AI without disclosing its use?

There are some reports of researchers who experienced this (Wolkovich 2024). As specified in question 18, there are, at present, no software tools that can reliably determine AI-generated text. Authors should ask for a detailed explanation of why their text or parts of it are suspected of being AI-generated. It is helpful to keep an internal record of the work and writing process so that the documentation can be used to trace or prove that AI was not used.

21. Are there any AI applications that can be recommended from the perspective of research integrity?

Which AI application is the most suited for a particular task depends on different factors. Researchers should choose an AI tool not only based on its features and overall performance, but should also check if their preferred tool complies with legal regulations. Data privacy and confidentiality play a key role here. Many AI tools are also known for their inherent weaknesses and risks (see question 17). Authors should be aware of these and carefully check any AI-generated output. Furthermore, also ethical questions arise in connection with certain AI applications and AI applications in general. The decision to use AI or one particular AI application should therefore be carefully considered at the start of each project. Resnik und Hosseini (2024) might constitute a useful guide for authors in that respect as it offers a detailed list how to transparently disclose the use of AI, which among others things, includes reflection questions on the functions and limits of AI tools.

22. What is the connection between copyright and AI-generated output?

Questions concerning the connection of copyright and AI tools may focus on the generated output (e.g. Do I own the copyright to the content I have generated with the help of AI?) as well as the input uploaded into an AI tool (Can I infringe on other people's copyright when working with AI?). These and related questions have been addressed in Roman Konertz 2023 and Ulrike Verch 2024 (both in German).

23. How can I prevent potential conflicts resulting from working with AI?

In the current dynamic situation with many researchers starting to incorporate AI tools into their work while recommendations and guidelines are still evolving, researchers should familiarise themselves at the outset with the AI guidelines that apply to them and correctly assess their own expertise in relation to AI. Researchers working in teams or in long-term projects with fluctuating staff should facilitate an open dialogue on AI use and create a transparent internal documentation so that everyone involved knows if and how AI was used. This includes research outputs apart from manuscripts such as software code, presentation slides and scripts, posters and more.

24. Which AI guidelines are relevant for me?

This depends on which AI tools you are using for what purpose. If you use AI tools for writing manuscripts, you should follow institutional guidelines (if present) as well as publishers' editorial policies. Give preference to guidelines from within your field, if they exist, especially if they contain criteria for disclosing the use of AI that are stricter than those found in editorial policies. Regarding PhD dissertations examination regulations apply. Furthermore, the topic should be addressed with the supervisors. For grant applications see question 12. Any potential contradictions, e.g. between institutional guidelines and editorial policies, should be communicated at an early stage.

25. Why are there no recommendations on AI by the German Research Ombudsman?

The German Research Ombudsman addressed the issue of recommendations in 2023 by convening an expert panel to discuss the relation between AI and RI. The results of that workshop were written up in a short report, which was the basis for a longer article on the topic, published in the December 2023 issue of the *Zeitschrift für Bibliothekswesen und Bibliographie* (ZfBB) (both only available in German). One key result of the workshop was that participants agreed that most questions concerning the RI-compliant use of AI are either already implicitly addressed by the DFG Code of Conduct or need to be discussed within individual fields in order to do justice to field specific criteria. Thus a general set of recommendations was not prepared. The German Research Ombudsman can advise others on drafting recommendations.

References

All references are listed below, but are also directly linked in the text. Additionally, you can find a collection of sources on AI and RI on our website.

Literature

Frisch, Katrin, Felix Hagenström und Nele Reeg. Textgenerierende KI und gute wissenschaftliche Praxis. *Zeitschrift für Bibliothekswesen und Bibliographie* 70, Nr. 6 (15. Dezember 2023): 326–36. https://doi.org/10.3196/186429502370667.

Gao, Catherine A., Frederick M. Howard, Nikolay S. Markov et al. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers, *npj Digital Medicine*, Volume 6, Issue 75. https://doi.org/10.1038/s41746-023-00819-6

Henderson, Peter, Xuechen Li, Dan Jurafsky, et al. (2023). Foundation Models and Fair Use [Preprint]. *Arxiv*. https://doi.org/10.48550/arXiv.2303.15715

Hosseini, Mohammad, David B. Resnik und Katie Holmes (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics*, Volume 19, Issue 4. https://doi.org/10.1177/17470161231180449

Hosseini, Mohammad, Serge P.J.M. Horbach (2023). Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res Integr Peer Rev* 8, 7. https://doi.org/10.1186/s41073-023-00133-5

Konertz, Roman (2023). Urheberrechtliche Fragen der Textgenerierung durch Künstliche Intelligenz: Insbesondere Schöpfungen und Rechtsverletzungen durch GPT und ChatGPT. *Wettbewerb in Recht und Praxis*, Jahrgang 69. https://doi.org/10.18445/20230613-111651-0

Oertner, Monika (2024). ChatGPT als Recherchetool?: Fehlertypologie, technische Ursachenanalyse und hochschuldidaktische Implikationen. *Bibliotheksdienst* 58, no. 5 : 259-297. https://doi.org/10.1515/bd-2024-0042

Pope, Audrey (2024). NYT v. OpenAI: The Times's About-Face. Harvard Law Review. https://harvardlawreview.org/blog/2024/04/nyt-v-openai-the-timess-about-face/

Resnik, David B., Hosseini, Mohammad (2024). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. *AI Ethics*. https://doi.org/10.1007/s43681-024-00493-8

Seadle, Michael (2017): *Quantifying Research Integrity*. San Rafael, California: Morgan & Claypool.

Verch, Ulrike (2024). Per Prompt zum Plagiat? Rechtssicheres Publizieren von KI-generierten Inhalten. *API*, Band 5, Nr. 1. https://doi.org/10.15460/apimagazin.2024.5.1.

Weber-Wulff, Debora, Alla Anohina-Naumeca, Sonja Bjelobaba, et al. (2023): Testing of detection tools for AI-generated text. *International Journal for Education Integrity*, Volume 19, Issue 26. https://doi.org/10.1007/s40979-023-00146-z

Wolkovich, E.M. (2024). 'Obviously ChatGPT' — how reviewers accused me of scientific fraud. *Nature*; Career Column. https://doi.org/10.1038/d41586-024-00349-5

Policies, Guidelines, Recommendations

ACS Nano: Best Practices for Using AI When Writing Scientific Manuscripts, 2023, 17, 5, 4091–4093. https://doi.org/10.1021/acsnano.3c01544

American Psychological Association: How to Cite ChatGPT: https://apastyle.apa.org/blog/how-to-cite-chatgpt

Berlin Universities Publishing: 'Guideline for Dealing with Artificial Intelligence' and 'Handout for Citing AI Tools', 2024 https://www.berlin-universities-publishing.de/en/ueber-uns/policies/ki-leitlinie/index.html

Chicago Manual of Style: Citing content developed or generated by artificial intelligence: https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html

Deutsche Forschungsgemeinschaft: Guidelines for Safeguarding Good Research Practice, 2019: https://wissenschaftliche-integritaet.de/en

Deutsche Forschungsgemeinschaft: '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', 2023: https://wissenschaftliche-integritaet.de/en/use-generative-models/

Elsevier: Generative AI policies for journals: https://www.elsevier.com/de-de/about/policies-and-standards/generative-ai-policies-for-journals

EU AI Act (2024/1689): https://artificialintelligenceact.eu/the-act/

European Commission: Living guidelines on the responsible use of generative AI in research, 2024: https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf

Frontiers: Artificial intelligence: fair use and disclosure policy: https://www.frontiersin.org/guidelines/policies-and-publication-ethics

ICMJE: Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals, 2024: https://www.icmje.org/icmje-recommendations.pdf

Modern Language Association: How do I cite generative AI in MLA style?: https://style.mla.org/citing-generative-ai/

Science Journals: Editorial Policies: https://www.science.org/content/page/science-journals-editorial-policies#image-and-text-integrity

Springer Nature: Editorial Policies: Artificial Intelligence: https://www.springernature.com/gp/policies/editorial-policies

Taylor & Francis: AI Policy: https://taylorandfrancis.com/our-policies/ai-policy/?_ga=2.202881233.1550617214.1729762987-990921615.1729762987

WAME: Recommendations on Chatbots and Generative Artificial Intelligence in Relation to Scholarly Publications, 2023: https://wame.org/page3.php?id=106

Wiley: Best Practice Guidelines on Research Integrity and Publishing Ethics: Artificial Intelligence: https://authorservices.wiley.com/ethics-guidelines/index.html#22