Katrin Frisch
Discussion Hubs to
Foster Research Integrity
ombudsgremium.de

FAQ
Artificial Intelligence
&
Research Integrity

Version 2

This FAQ compiles a set of questions that we frequently receive regarding artificial intelligence (AI) and research integrity (RI). This version 2 is an updated and extended version of the original FAQ document, published in November 2024. The scope of the FAQ is limited to generative AI, especially Large Language Models (LLMs) that are the basis of many tools available to researchers since the launch of ChatGPT in November 2022. The aim is to offer guidance in a fast-paced debate without being prescriptive. This FAQ does not constitute an official statement on the use of AI by the Ombuds Committee for Research Integrity in Germany (OWID). Instead, it describes the status quo, contextualises existing recommendations, identifies gaps and provides further literature. The intended audience are researchers as well as research integrity advisors. 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 (Selbstständigkeitserklärungen) as well as individual decisions made by lecturers.

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 still need to be defined.

The following overview provides a summary of AI policies of the major publishers and publishing associations (last update: October 2025)

	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	not allowed	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 further specified	highly restricted	https://www.icmje.org/icmje-recommendations.pdf
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
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
Taylor & Francis	AI ≠ author	Use of AI needs to be declared in manuscript	not allowed	highly restricted	https://taylorandfrancis.com/our-policies/ai-policy/
Wiley	AI ≠ author	Use of AI needs to be declared in manuscript	not specified (journal policy) restricted (book policy)	highly restricted	Journal authors: https://authorservices.wiley.com/ethics-guidelines/index.html#22 Book authors: https://www.wiley.com/en-us/publish/book/resources/ai-guidelines/
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.

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. 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

Some policies offer detailed information on how to declare the use of generative AI. For example, Wiley's AI Policy for book authors lists which use cases need to be declared and which aspects need to be included in the declaration. Authors can also find sample declarations. Smaller publishing houses, such as Berlin Universities Publishing (BUP), have also developed more detailed recommendations. A number of publishing houses expect authors to include a (critical) reflection on their AI use in the declaration (e.g. Wiley, PLOS one). Researchers, who have written on the topic, further suggest specifying the member of the team that made use of AI (see Hosseini et al. 2023) and propose detailed reflections on the technical functionalities and limitations of the AI used (Resnik and Hosseini 2024). The latter exceeds common information required for disclosure, however it may be useful to reflect on the suitability of an AI application as well as its inherent weaknesses and limitations, which requires authors to familiarise themselves with the tool in question.

Other approaches to declare the use of generative AI are the Artificial Intelligence Disclosure (AID) Framework and the AI Attribution Toolkit. The Artificial Intelligence Disclosure (AID) Framework by Kari D. Weaver is inspired by the Credit Roles Taxonomy (CRediT). It recommends listing the 'AI tools used and the manner in which they were used' throughout the research process in a brief statement appended to the manuscript. The AI Attribution Toolkit – developed by researchers at IBM Research and based on He et al. 2025 – takes inspiration from the Creative Commons licencing system and focuses on disclosing the proportion of AI use in the form of a concise attribution statement.

In general, the disclosure 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 8-11).

4. Which tools are covered by the term AI?

No single definition exists for the term Artificial Intelligence and the term itself has been subject to debate. In the general debate ChatGPT has become a synonym for generative AI, especially Large Language Models (LLMs). 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. A clear demarcation between AI applications (that require disclosure when used) and other software tools (that do not require disclosure) might become increasingly difficult. Guidelines for researchers mostly focus on generative AI, especially LLMs. AI applications that are only used to improve language use and style fall into a grey area (see question 9). AI applications that can assist in the first steps of a research process (for example, 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 use cases are classified as potentially 'substantial use' and thus need to be declared along with the AI tool used. 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.

5. 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. Some policies propose to disclose AI use in a separate statement at the end of a manuscript. Which part of the manuscript is the most suitable may also depend on field-specific standards as well as the length and 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).

6. 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), the Chicago Style Manual, and the Modern Language Association (MLA). Please ensure that the citation style you use includes all the mandatory information requested by your institution/journal.

7. Do prompts need to be disclosed?

At present there is no consensus on this question. Of the major 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 state that prompts can, but do not necessarily need to be included, offering different suggestions on where they could be provided. The APA specifies that researchers should document the prompts 'for their own records'. Among researchers disclosure of prompts is a contentious issue: some question the usefulness of listing prompts as it neither corresponds to the actual (often iterative) use of AI nor does it necessarily offer increased transparency, since 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 serves 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. Additional transparency concerning the use of generative AI can also be provided by other means, such as reflection statements (see question 3).

8. 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. Wiley's AI Policy for book authors is one of the few exceptions, specifying that generating and modifying code with the help of generative AI requires disclosure. 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, which is especially important if the code is made available to others.

9. 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 versus tools that check grammar and spelling, like Grammarly). AI policies often do not cover the latter, i.e. their use does not need to be disclosed (see the AI policies by Elsevier or Wiley). Wiley, for example, specifies that '[b]asic grammar or spell checking' as well as '[s]imple language polishing' do not require declaration. The DFG, likewise, notes 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.' This sentiment is echoed by researchers who suggest that declaring the use of generative AI as a writing aid should be voluntary (Hosseini et al. 2025).

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

10. 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. There are some exceptions, which differ in their assessment of the matter. Wiley's AI Policy for book authors requires declaration when generative AI is used for translations. On the other hand, the DFG guideline on AI lists translation tools among those that do not require documentation (see also question 9). Likewise the BUP Guideline for Dealing with AI classifies translation 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' (VdÜ / A*ds / IGÜ – Offener Brief zur KI-Verordnung, translated by K.F.). In disciplines like Modern Languages or when translating literary texts documenting the use of translation tools is recommended.

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

Depending on how 'inspiration' is defined, some recommendations on that matter can be found. For example, the BUP Guidelines for Dealing with AI suggest that authors should add a general note or disclose the use of AI for inspiration in the methods section. The Living Guidelines on the Responsible Use of Generative AI in Research issued by the European Commission state that AI tools that 'have been used substantially in [the] research process[…] should be made transparent', which may include 'identifying research gaps' and 'developing hypotheses' (page 7-8). 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 and extent of the AI-generated ideas for the publication in question.

12. Can I use AI for writing grant proposals?

Please refer to the information provided by the respective research funding organisation. The 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 generating 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). The most detailed AI Policy on that issue is at present Wiley's AI Policy for book authors. It lists both acceptable use cases for AI-generated images as well as prohibited use cases. Furthermore, it states which requirements permitted use cases (which include flow charts and visualisations) need to fulfil. The use of AI-generated images that pretend to be factual/evidential images like Western Blots, specimens, samples or artefacts is prohibited. From the perspective of RI they constitute an act of deception and could be considered data fabrication (and thus research misconduct).

No guidelines yet exist 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 must not be used to feign genuine research data or results. 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 processes (like diagrams or flow charts) the use of AI could be permissible, potentially requiring disclosure. As with AI-generated text it is the responsibility of the researchers involved to check the results for accuracy. In addition to the general question whether the use of AI-generated images is permissible in presentations and posters, there are at the moment no clear guidelines on disclosure in this context. Scholars can consult existing guidelines for declaring AI use in publications. Due to its concise form, the AI Attribution Toolkit might be especially useful in this context (see question 3).

14. Can I use AI for peer review?

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 protection, 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', which should not be outsourced to an AI tool. (Hosseini und Horbach 2023; see also Bergstrom and Bak-Coleman 2025). The use of generative AI in peer review is still subject to debate. According to a recent study by Nature, the majority of respondents do not consider it appropriate to use an AI tool to review a manuscript. However, an AI tool that assists with peer review, for example by answering questions about a manuscript, was seen as potentially appropriate by more than 50% of respondents. Although the use of generative AI to review manuscripts is currently not permitted by publishers' policies, studies show that a small minority of reviewers breach AI policies. Publishing houses themselves make use of various AI tools to pre-screen submitted manuscripts.

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 at the start of each project whether all authors agree on the use of AI tools and the extent of their use. This is especially relevant in trans- and interdisciplinary teams, since there is a greater likelihood of conflicting views on aspects such as text production (see also question 9). Teams can benefit from good internal documentation on the use of AI. In case of conflicts or breaches of research integrity, good documentation allows others to trace the work process and genesis of the research manuscript. Authors publishing in teams may also consider following the suggestions of 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, having been trained on huge datasets, inadvertent plagiarism is not one of the common risks of AI. This is based on an understanding of plagiarism as the unattributed reuse of contributions by third parties. This definition presupposes that there is an identifiable source or person from whom text has been taken verbatim or paraphrased without attribution. There are examples of well-known AI applications that can reproduce certain texts (almost) verbatim, despite their stochastic mode of operation (see Cooper et al. 2025, Henderson et al. 2023 or the case of the New York Times versus OpenAI, e.g. Pope 2024). However, in these studies specific prompting was used to achieve exactly these results.

In some fields within the humanities where 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.

Another aspect that should be noted is that in examination regulations issued by universities another definition of plagiarism concerning student papers can be found. As Ulrike Verch notes: using AI-generated output 'without labelling and comprehensive revision could constitute plagiarism, as university examination regulations generally require that submitted work be produced independently without any aids other than those specified and permitted' (Verch 2023, page 11, translated by K.F.).

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 generative AI tools. These include missing or incorrect references, errors in direct quotes, fabricated quotes or references, outdated information as well as the reproduction of bias and prejudice. The high 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). Many generative AI tools now carry a disclaimer that the results could contain errors. Yet, researchers should also be wary of tools that claim to produce reliable results and verify all AI-generated output. 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.

At a more general level, generative AI tools carry further risks. These include epistemic risks (see, for example, Messeri and Crockett 2024; Schütze 2025), de-skilling, an increase in paper mill output, overburdening of research infrastructure by bots (e.g. Weinberg 2025), loss of digital autonomy (Bahr 2024), dependency on proprietary services and commercial actors, copyright violations as well as exploitation of human labourers and natural resources (Hao 2025). A comprehensive list of risks in connection with AI can be found at the MIT AI Risk Repository.

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 reached 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.

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, which is a breach of existing editorial policies/recommendations. Editorial teams should discuss how breaches of their AI policy should be handled. Some editorial policies have set out a basic procedure what to do in these cases. If there are less tangible 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 in such cases 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 making 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 also check if their preferred tool complies with legal regulations. Data privacy and confidentiality play a key role. 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.

The decision to work with AI or a specific AI tool should be made in advance with care. Before using an AI tool in their research, researchers should thoroughly study the tool in question. This includes ethical as well as integrity issues of that particular tool as well as AI tools in general (see also question 17). Moreover, looking at the technical aspects of AI tools (training data, model weights, model cards) can be helpful. However, for proprietary tools this information is not always available. To address this aspect, researchers can opt for (more) open models. The European Open Source AI Index offers a good overview. Resnik und Hosseini (2024) might constitute a useful guide for authors as it offers a detailed list on how to transparently disclose the use of AI, which among other things, includes reflection questions on the functions and limits of AI tools. Similarly, the MLA Guide to AI Literacy (although targeted at students) may be used as a helpful guide by researchers.

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

Questions concerning the connection between 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 by Roman Konertz (2023) and Ulrike Verch (2024) as well as in the FAQ of the German Publishers and Booksellers Association (all in German).

Questions concerning copyright also arise in connection with the training of LLMs. Whether this practice falls under the fair use principles of copyright is at the moment subject of several US court hearings. A first decision was reached in June 2025 in the court case Bartz/Graeber/Wallace versus Anthropic that addressed the use of copyrighted material that had been obtained from shadow libraries. The verdict saw the use of copyrighted material for the training of LLMs as covered under the fair use principles, but not the creation of a training corpus with the help of pirated material ([Bartz v. Anthropic PBC, 3:24-cv-05417, (N.D. Cal.)](#)).

23. Can I prevent my published texts from being used as training data?

In accordance with German copyright law (Urheberrecht) the agreement of rights holders is theoretically needed (see [FAQ of the German Publishers and Booksellers Association](#)). Researchers should check if the publisher with whom they intend to publish has set up a licensing agreement with an AI company. For Open Access texts with CC BY licences there are basically two options. The more open licences CC BY and CC BY SA allow the use for training purposes. The non-commercial CC licences only allow the use for training if all aspects of the training and the end result are not commercial. Material with ND licences cannot be used for training purposes. (see [Creative Commons: Using CC-Licensed Works for AI Training](#)). In practice copyrighted material (including pirated copies, see [question 22](#)) has repeatedly been used for training purposes without the permission of authors (see for example [Reisner 2025](#)).

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

In the current dynamic situation, where many are 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. For PhD-projects, supervisors and PhD candidates should (unless institutional guidelines on the matter exist) agree on rules for AI usage and declaration. 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 other research outputs such as software code, presentation slides and scripts, posters, and more.

25. 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 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.

26. Why are there no recommendations on AI by OWID?

The Ombuds Committee for Research Integrity in Germany (OWID) 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) (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 implicitly addressed by the DFG Code of Conduct and other RI regulations 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. OWID 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

Bahr, Amrei (2024). Bridges Without Foundation? Why the Use of AI Tools in Academia Needs to Build on Ethics First. *FILOZOFIA*, 79, No 5, pp. 486 – 500. https://doi.org/10.31577/filozofia.2024.79.5.2

Bergstrom, Carl T. and Joe Bak-Coleman (2025). AI, peer review and the human activity of science. *Nature*. https://doi.org/10.1038/d41586-025-01839-w

Cooper, A.F., Gokaslan, A., Cyphert, A.B., Sa, C.D., Lemley, M.A., Ho, D.E., and Liang, P. (2025). Extracting memorized pieces of (copyrighted) books from open-weight language models. ArXiv. https://doi.org/10.48550/arXiv.2505.12546

Frisch, Katrin, Felix Hagenström and 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

Hao, Karen (2025). *Empires of AI: Dreams and Nightmares in Sam Altman's OpenAI*. New Yok. Penguin Press.

He, Jessica, Stephanie Houde, and Justin D. Weisz (2025). Which Contributions Deserve Credit? Perceptions of Attribution in Human-AI Co-Creation. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25).* 540, 1–18. https://doi.org/10.1145/3706598.3713522

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

Hinchliff Pearson, Sarah (2025). Understanding CC Licenses and AI Training: A Legal Primer. https://creativecommons.org/2025/05/15/understanding-cc-licenses-and-ai-training-a-legal-primer/

Hosseini, Mohammad, Gordijn, Bert, Kaebnick, Gregory E., & Holmes, Katie (2025). Disclosing generative AI use for writing assistance should be voluntary. *Research Ethics*. https://doi.org/10.1177/17470161251345499

Hosseini, Mohammad, David B. Resnik and 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

Messeri, Lisa, Crockett, M.J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature* 627, 49–58. https://doi.org/10.1038/s41586-024-07146-0

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/

Reisner, Alex (2025). The Unbelievable Scale of AI's Pirated-Books Problem. *The Atlantic.* https://www.theatlantic.com/technology/archive/2025/03/libgen-meta-openai/682093/

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.

Schütze, Paul (2025). Zwischen Technologie und Ideologie: Ein Blick auf die un/realen Räume hinter KI. In: Clara Kindler-Mathôt/Didem Leblebici/Giacomo Marinsalta/Till Rückwart and Anna Zaglyadnova. *Un/Reale Interaktionsräume*. Bielefeld: transcript Verlag: 81-100. https://doi.org/10.14361/9783839471463-006

Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2024). The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks from Artificial Intelligence. https://doi.org/10.48550/arXiv.2408.12622

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.

Weaver, Kari D. (2024). The artificial intelligence disclosure (AID) framework: An introduction. *C&RL News*, 85(10), 407-411. https://crln.acrl.org/index.php/crlnews/article/view/26548

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

Weinberg, Michael (2025). Are AI Bots Knocking Cultural Heritage Offline? The GLAM-E Lab. https://www.glamelab.org/products/are-ai-bots-knocking-cultural-heritage-offline/

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/cite-generative-ai-references

AI Attribution Toolkit: https://aiattribution.github.io/

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

Börsenverein des Deutschen Buchhandels: Künstliche Intelligenz im Verlagsbereich: FAQ zu generativer KI. https://www.boersenverein.de/beratung-service/recht/kuenstliche-intelligenz/

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, V2, 2025: 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-updated-revised/

PLOS One: Ethical Publishing Practice/ Artificial Intelligence Tools and Technologies: https://journals.plos.org/plosone/s/ethical-publishing-practice#loc-artificial-intelligence-tools-and-technologies

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/

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

Wiley: AI Guidelines [for book authors]: https://www.wiley.com/en-us/publish/book/resources/ai-guidelines/

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