

Living guidelines on the **responsible use of generative AI in research**

ERA Forum
Stakeholders'
document

Fourth version, October 2026



Living guidelines on the responsible use of generative AI in research

European Commission
Directorate-General for Research and Innovation
Directorate E-Prosperity
Unit E4 - AI in Science & Critical Technologies

To receive news and event alerts on AI in science, sign up [here](#)

Email eu-ai-in-science@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu

European Commission
B-1049 Brussels

Manuscript completed in October 2026
Fourth edition

Neither the European Commission nor any person acting on behalf of the European Commission is responsible for the use that might be made of the current information.

Luxembourg: Publications Office of the European Union, 2026

© European Union, 2026



The Commission's reuse policy is implemented under Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39, ELI: <http://data.europa.eu/eli/dec/2011/833/oj>).

Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed, provided appropriate credit is given and any changes are indicated.

For any use or reproduction of elements that are not owned by the European Union, permission may need to be sought directly from the respective rightholders. The European Union does not own the copyright in relation to the following elements:

Cover: © MicroOne #288703015, creativeteam #323412491, skypicsstudio #286372753, Viktoriia #345410470 #479686001, 2020. © Tartila #279519357, © naum #481091871, 2022. © AKrasov #473038224, 2024. Source: Stock.Adobe.com

Living guidelines on the
**responsible use of
generative AI in research**

ERA Forum Stakeholders' document

Table of contents

1. INTRODUCTION 3

2. RECOMMENDATIONS FOR RESEARCHERS..... 5

3. RECOMMENDATIONS FOR RESEARCH ORGANISATIONS..... 9

4. RECOMMENDATIONS FOR RESEARCH FUNDING
ORGANISATIONS 10

5. ONGOING DEVELOPMENTS OF GENERATIVE AI IN RESEARCH
..... 12

6. BACKGROUND AND POLICY CONTEXT 15

ABOUT THESE LIVING GUIDELINES 15

RESEARCH INTEGRITY 16

TRUSTWORTHY AI.....17

REFERENCES 18

ARTICLES AND PAPERS.....18

GUIDELINES ANALYSED 19

1. INTRODUCTION

Artificial intelligence (AI) has become an integral part of our daily lives with disruptive and overwhelming acceleration of AI capabilities. The developments under the general category of generative AI¹ have become the most popular and widely adopted, generating different types of outputs (text, code, data, images, music, voice, videos, etc.), based on instructions provided by the user (like prompts). The capabilities of generative AI models are evolving significantly, in particular in reasoning and planning capabilities. This has enabled the development of agentic AI², (for example in coding activities), that can tackle complex problems and tasks, with some degree of planning and autonomy.

Generative AI in research

Research is one of the sectors that is being most impacted by generative AI. AI has great potential for accelerating scientific discovery. Positive and simple examples of the use of these tools by researchers include supporting non-native speakers in producing texts in other languages, producing quickly text summaries from large and different sources of information, automatically retrieving and contextualising a wide body of knowledge and helping to write code to analyse data³. More advanced uses are emerging in which AI could be a fundamental partner or tool in the research process, as seen in recent mathematical problems solved with, or by generative and agentic AI.

The technology also entails the risk of misuse. Some risks are due to the tool's technical limitations, and others have to do with the (intentional or unintentional) use of the tool in ways that erode sound research practices. Additional risks for research in Europe could stem from the proprietary nature of some of the tools (for example, lack of openness, fees to access the service, use of input data), the concentration of ownership or the undesirable transfer of critical technology and intellectual property⁴.

The impact of generative AI on research and various aspects of the scientific process calls for reflection. While the potential of this technology is enormous for science, in some respects, using these tools could challenge research integrity and raise questions about their contribution to deceptive scientific practices and misinformation. Researchers and organisations should therefore consider what could go wrong, what is at stake, and what safeguards are appropriate, while supporting its beneficial adoption.

Origins of these guidelines

These Living Guidelines were developed within the European Research Area Forum⁵ as a common reference for the responsible use of generative AI in research. They were initiated

¹ "Generative AI (or GenAI) refers to AI systems that can create new content like text, images, music, code, or video. These systems learn patterns from training data and generate novel outputs that resemble the original data (...). These models power applications including chatbots, code generation, and creative tools." <https://hai.stanford.edu/ai-definitions/what-is-generative-ai>

² "AI agents" generally refer to systems that can perceive and act upon their environment with a degree of autonomy, using tools as needed to achieve specific goals and adapt to changing inputs and contexts. OECD (2026), "The agentic AI landscape and its conceptual foundations", OECD Artificial Intelligence Papers, No. 56, OECD Publishing, Paris, <https://doi.org/10.1787/396cf758-en>.

³ Results of a survey by Nature, <https://doi.org/10.1038/d41586-023-02980-0> and follow up survey <https://doi.org/10.1038/d41586-024-03940-y>

⁴ Council Recommendation of 23 May 2024 on enhancing research security ([ST/9097/2024/INIT](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R0907))

⁵ The [ERA Forum](https://eraforum.eu/) is a platform for dialogue and cooperation between national governments, the European Commission, and stakeholders in research and innovation

in response to the rapid proliferation of AI guidance from different institutions and sectors, bringing a voice of authority in a complex landscape.

One of the goals of these guidelines is that the scientific community uses this technology in a responsible manner, embracing the positive uses while being aware of its risks and limitations. They are intended to support researchers, research organisations and research funding organisations in applying existing principles of research integrity and responsible AI use in a rapidly changing research environment. Care has been taken to allow for appropriate adaptation to different research disciplines, institutions, research activities and national contexts.

Yet, the development of a robust framework for generative AI in scientific research cannot be the sole responsibility of policymakers at European and national levels. Universities, research organisations, funding bodies, research libraries, learned societies, publishers, research managers, research support staff and researchers at all stages of their careers are essential in shaping the discussion on AI and how it can serve the public interest in research.

How to read these guidelines

These guidelines are non-binding. They complement, but do not replace, applicable legislation and relevant institutional, disciplinary, funder and other requirements. Where specific legal or policy requirements apply, these take precedence. They take into account key principles of research integrity⁶.

These guidelines also invite the whole scientific community to reflect critically on when and how AI use is needed. Institutional users of these guidelines are encouraged to adapt them to their specific contexts and situations, keeping proportionality in mind.

The recommendations in the next chapters are not new requirements, or separate standards for AI-assisted research. They are a translation of current expectations, applying the underlying principles into a research ecosystem where AI is increasingly integrated. Understanding these principles helps to navigate the recommendations. **AI does not change the principles of responsible research.**

- Always consider whether AI is the right and best tool for a particular task
- Accountability remains with researchers/users and the institutions with which they are affiliated
- Confidentiality is non-negotiable
- Transparency and disclosure are necessary to build and maintain trust

⁶ More details in the background section

2. RECOMMENDATIONS FOR RESEARCHERS

For generative AI to be used in a responsible manner, researchers should:

1. Understand that they remain ultimately responsible for scientific output.

- Researchers are accountable for the integrity of the content⁷ they produce generated by or with the support of AI tools.
- Researchers maintain a critical approach to using the output produced by generative AI and are aware of the tools' limitations, such as bias⁸, hallucinations⁹ and inaccuracies. They ensure appropriate oversight and verification of AI contributions.
- AI systems are neither authors nor co-authors nor owners. Authorship implies agency and responsibility, so it lies with researchers.
- When aiming to implement complex tasks autonomously, for example with AI agents, researchers identify and ensure relevant check points throughout the process to ensure that critical decisions (for example methodological) are validated by the researchers.
- The use of AI agents may never be considered a delegation of control or responsibility. Researchers ensure that they remain accountable and apply sufficient human oversight and verification in all phases of any scientific endeavour that uses any form of delegation to AI agents.

2. Use generative AI transparently.

- Researchers disclose which AI tools or systems were used substantially¹⁰ in their research or reporting processes and describe what they contributed or were used for. In doing so, they follow the guidelines of their discipline, journal (for publications), funder (for proposals and evaluation reports) or institution.
- Researchers also describe meaningful information about oversight and verification strategies (such as measures to identify hallucinations, mitigate bias, etc.) in order to enable readers to understand and assess the validity of AI-assisted work.
- Researchers consider the stochastic or otherwise non-deterministic nature of generative and agentic AI, which may produce different outputs from the same inputs. They assess how this may affect the reproducibility and robustness of their work. Authors decide whether prompts, logs, model versions, parameters and other information should be retained or shared, based on whether these are needed to understand, reproduce or audit the research, or to assess its validity. This is particularly relevant when AI agents perform sequences of research tasks or contribute to methodological decisions.

3. Pay particular attention to issues related to privacy, confidentiality and intellectual property rights when sharing sensitive or protected information with AI tools.

- Researchers remain mindful that generated or uploaded input (text, data, prompts, images, etc.) could be used for other purposes, such as the training of AI models. Therefore, they protect unpublished or sensitive work (such as their own or others'

⁷ Content should not be fabricated, falsified or plagiarised. For example, researchers do not use generative AI to falsify, alter or manipulate original research data. More details in Section 3.1 of [The European Code of Conduct for Research Integrity](#), ALLEA 2023.

⁸ Ethical prompting should be applied. See the [Ethical Prompting for Generative AI guidelines of the ETH Zurich](#) for more information. On sycophantic behaviour for example [this paper](#).

⁹ "hallucination" refers to the generation of false, nonsensical, or inaccurate information by large language models (LLMs) or other generative AI systems.

¹⁰ In this case, for example, using generative AI as a basic editorial support tool is not a substantial use. However, using it as a method or source, interpreting data analysis, carrying out a literature review, identifying research gaps, formulating research aims, developing hypotheses, etc. is substantial use.

unpublished work) by taking care not to upload it into an external AI system unless there are assurances that the data will not be re-used, e.g., to train future language models or to the untraceable and unverifiable reuse of data.

- Researchers take care not to provide third parties' personal data to external generative AI systems unless the data subject (individual) has given their consent and researchers have a clear goal for which the personal data are to be used so compliance with EU data protection rules¹¹ is ensured¹².
- Researchers understand the technical, ethical and security implications regarding privacy, confidentiality, intellectual property rights and research security. They check, for example, their institutional guidelines, the privacy options of the tools, who is owning or managing the tool (public or private institutions, companies, etc.), where the tool is running and implications for any information uploaded¹³. This could range from closed environments, hosting on a third-party infrastructure with guaranteed privacy, to open internet-accessible platforms.
- If an AI agent is granted access to sensitive information following a suitability assessment, the researcher shall instruct and constrain potential confidentiality breaches and remain accountable for how the agent uses that sensitive information.

4. Respect applicable national, EU and international legislation. In particular, the output produced by generative AI can be especially sensitive in relation to the protection of intellectual property rights and personal data.

- Researchers pay attention to the potential for plagiarism (text, code, images, etc.) when using outputs from generative AI. Researchers respect others' authorship and cite their work where appropriate. The output of a generative AI (such as a large language model) may be based on someone else's results and require proper recognition and citation¹⁴.
- The output produced by generative AI can contain personal data. If this becomes apparent, researchers are responsible for handling any personal data output responsibly and appropriately, and EU data protection rules are to be followed.
- Researchers are responsible for taking steps to ensure that AI agents do not breach any applicable legislation in their efforts to reach their goals. Researchers should be mindful that they are not exempted from their legal responsibilities because an agent acted on their behalf.

5. Continuously learn how to use generative AI tools properly to maximise their benefits, including by undertaking training.

- Generative AI tools are evolving quickly, and new ways to use them are regularly discovered. Researchers stay up to date on the best practices and share them with colleagues and other stakeholders.
- Researchers aim to minimise the environmental impact¹⁵ of generative AI, by evaluating whether and which AI tool is best suited for the intended task and by using the most effective prompting techniques.

¹¹ https://commission.europa.eu/law/law-topic/data-protection/eu-data-protection-rules_en. See also the [First Orientations for ensuring data protection compliance when using Generative AI systems](#) by the European Data Protection Supervisor.

¹² Apart from personal consent there could be other lawful basis that could justify the processing of personal data.

¹³ This includes research-security elements (e.g. sensitive/dual-use research data being exposed to non-trusted foreign entities).

¹⁴ The technology is evolving, and citations quality will improve, but the final responsibility on the citation and its correctness remains with the researcher.

¹⁵ This can be seen as part of the broader initiatives to reduce the carbon footprint of research, following the principles for sustainable coding and sustainable research (see for example <https://doi.org/10.1371/journal.pcbi.1009324> and <https://doi.org/10.1371/journal.pcbi.1008148>). Preliminary data on the environmental cost of generative AI can be found in the [AI Index Report 2023](#).

6. Assess critically any potential use of generative AI when evaluating funding proposals or undertaking peer review of research outputs. Evaluators, peer reviewers and editors should:

- Comply with the rules and guidance applicable to the specific evaluation or review process. Follow relevant institutional policies and any restrictions or prohibitions on AI use, namely those from funding organisations or journals and publishers.
- Ensure that the judgement task is not delegated to an AI tool when assessing scientific quality and merit of proposals and papers. Remain fully accountable for all assessments, recommendations, judgements and decisions.
- Use generative AI only for certain supporting tasks¹⁶. Use it only when permitted and where it is the most appropriate tool available.
- Remain aware of and address the challenges associated with generative AI, including:
 - Confidentiality and data protection: do not disclose unpublished proposals, manuscripts, review reports or key ideas through non-secured AI tools, including through excerpts or prompts. Ensure the use of secure, trusted tools¹⁷; preferably provided by the institution, funder or publisher.
 - Bias and overreliance: AI models can introduce or amplify bias. Verify factual claims, references and any substantive information used against original sources.
- Be transparent about its use. Disclose substantive AI use, and any other use required by applicable policies, in line with relevant community and involved institutions standards.

Generative AI and bias

Generative AI models can exhibit limitations that impact both their responses and usability. The most common types are:

- **Training Data Bias:** Biases in the data used to train generative AI models can lead to skewed responses, reflecting inaccuracies or systemic biases within the source material.
- **Prompt Bias:** Models may align their answers with the perceived beliefs or preferences of the user, a phenomenon known as sycophantic behaviour, potentially leading to misleadingly agreeable or biased outputs.
- **Invented Citations and Incorrect Summaries:** Generative AI models may generate plausible sounding but incorrect citations, which can mislead users who rely on those sources for research or verification. Even when the citations are correct (the paper exists, has the proper title...), the AI-generated summary of those papers may be incorrect. Therefore, researchers are responsible for checking all the references and content of the citations and the summaries.
- **Interpretability:** generative AI models operate as "black boxes," making it difficult to understand how specific responses are generated. This opacity underscores the importance of cross validation, especially in automated data analysis, where model responses can significantly impact conclusions.

¹⁶ For example, it can be used to clarify concepts, background research or support language-polishing. But AI can't be used to draft evaluation reports or assess scientific merit. More detailed guidance on the AI in review of proposals can be found in for example in the [EC corporate guidance for experts](#) or the [ERC Q&A](#) as general guidance.

¹⁷ E.g., do not rely on public instances of AI systems and prefer bespoke solutions with clear data management and governance in place.

Disclosure of AI use

In principle, AI use should be documented and reported like any other important research method or tool. This means that its use is best disclosed when it has (or could have) a substantive or otherwise relevant impact on the research process or its reporting. In particular, disclosure is expected if the use of AI shapes or influences human judgement, or if the contribution of AI may affect the results, their meaning or interpretation. That could likely be the case when AI takes a role in literature review, identifying research gaps, developing hypotheses, data analysis or interpretation, etc. Purely technical or minor uses that should not affect substance, such as spellchecking, grammar correction, polishing or formatting, generally do not need to be disclosed. In case of doubt, disclosure is advised.

Consider the use of generative AI by third parties that could be interacting with you/your organisation

- Researchers and organisations remain mindful that third parties may use generative AI tools. For example, in the context of meetings, note-taking, summarising discussions or creating overviews of documents (that could be confidential). Therefore, they can face similar risks to the ones described in this guidance, namely on confidentiality, data protection or intellectual property rights related to protected or original information.
- Inform third parties about the use of these kinds of tools and the potential risks, for example by sharing this guidance.
- In your contacts with third parties inform them if you are the one using those AI tools and how you are protecting collected information.

3. RECOMMENDATIONS FOR RESEARCH ORGANISATIONS

For generative AI to be used in a responsible manner, research organisations should:

- 1. Promote, guide and support the responsible use of generative AI in research activities.**
 - Research organisations provide and/or facilitate training for all career levels and disciplines, including for research managers and research support staff, on using generative AI, especially (but not exclusively) on verifying output, maintaining privacy, addressing biases and protecting intellectual property rights and sensitive knowledge.
 - Research organisations promote an atmosphere of trust where researchers are encouraged to transparently disclose the use of generative AI without concerns for adverse effects.¹⁸
 - Research organisations provide support and guidelines to ensure compliance with ethical and legal requirements (EU data protection rules, protection of intellectual property rights, etc.).
- 2. Actively keep track of the evolution and use of generative AI systems within their organisations.**
 - Research organisations remain mindful of the research activities and processes for which they use generative AI to better support its future use. This knowledge can:
 - be used to provide further guidance on using generative AI, help identify training needs and understand what kind of support could be most beneficial;
 - help anticipate and guard against possible misuse and abuse of AI tools;
 - be published and shared with the scientific community.
 - Research organisations analyse the limitations of the technology and tools and provide feedback and recommendations to their researchers.
 - Research organisations keep track of the environmental impact of generative AI within their organisations and promote awareness raising initiatives to help their staff pick the most sustainable option.
- 3. Reference or integrate these generative AI guidelines into their general research guidelines for good research practices and ethics.**
 - Using these guidelines as a basis for discussion, research organisations openly consult their research staff and stakeholders on the use of generative AI and related policies.
 - Research organisations apply these guidelines whenever possible. If needed, they could be complemented with specific additional recommendations and/or exceptions that should be published for transparency.
- 4. Whenever possible and necessary, implement locally hosted or cloud-based generative AI tools that they govern themselves¹⁹. This enables their employees to feed their scientific data into a tool that ensures data protection, confidentiality and research security**

¹⁸ Disclosure of the use of generative AI for assistance when writing may lead to a lower quality score during assessment, see for example Li *et al*, 2024, DOI: [10.18653/v1/2024.emnlp-main.279](https://doi.org/10.18653/v1/2024.emnlp-main.279) or Suchikova, Y., & Tsybuliak, N. (2025). The Purity Myth: Why Stigmatizing GAI in Academic Writing Is Harmful. *Science Communication*, 47(5), 753-761. <https://doi.org/10.1177/10755470241313233>.

¹⁹ Govern in this case means to establish and maintain the organisational rules, responsibilities and controls that ensure how it is developed, deployed and used safely. In this case can also be governed by trustworthy third parties that meet EU data protection and research security requirements,.

- Organisations ensure the appropriate level of cybersecurity of these systems, especially those connected to the internet.

4. RECOMMENDATIONS FOR RESEARCH FUNDING ORGANISATIONS

Research funding organisations operate in different contexts and follow different mandates and regulations, which may not align with a single set of guidelines. The recommendations below outline a set of measures and good practices for organisations to implement in ways that best fit their specific circumstances and objectives. For generative AI to be used in a responsible manner, research funding organisations should:

- 1. Promote and support the responsible use of generative AI in research.**
 - Research funding organisations design funding instruments that are open, receptive and supportive of the responsible and ethical use of generative AI technologies in research activities.
 - Research funding organisations require funded research and grantees to be in line with existing national, EU and international legislation (where applicable) and good practices for the use of generative AI.
 - Research funding organisations encourage researchers, research managers, research support staff and research organisations to use generative AI ethically and responsibly, including respecting legal and research standards requirements.
- 2. Analyse regularly the use of generative AI in internal processes, in particular during the evaluation of proposals.**
 - Research funding organisations remain fully responsible for the use of generative AI and agentic AI in their activities, in line with the accountability principle that emphasises responsibility and human oversight.
 - Generative AI could be used to improve internal processes for funders, keeping a balance between efficiency, human involvement, supervision and safeguards (including safeguards against biased, discriminatory or erroneous outcomes, particularly in screening, ranking, prioritisation or exclusion) and being transparent about its use.
 - When choosing and using generative AI tools, research funding organisations will carefully consider the tool's adherence to standards and challenges associated with its use and deployment²⁰. This is especially sensitive concerning the evaluation and funding of proposals, where confidentiality and non-delegation of the evaluation of scientific merit must be present throughout the whole funding process.
- 3. Guide evaluators on the use of generative AI in proposal evaluation processes, with a focus on data protection, confidentiality, transparency and the non-delegation of evaluative judgement.**
 - Research funding organisations provide guidance, similar to these guidelines, to ensure proper understanding by the evaluators of risks, challenges and limitations when using AI to support their evaluation activities. Guidance should be straightforward and proportionate to the associated risks.
 - If research funding organisations allow the use of generative AI in evaluation-related activities, they must provide active support for allowed uses and facilitate evaluators to disclose when and for what purposes AI has been used.

²⁰ Organisations especially consider elements concerning quality, transparency, integrity, data protection, confidentiality and respect for intellectual property rights, as well as its environmental impact or compliance with regulation like the AI Act.

- When research funding organisations provide access to internal AI tools to evaluators, they will provide appropriate details on when and how to use them, and on potential risks and limitations. They will also ensure that the tools respect confidentiality and data processing requirements according to relevant legislation and policies.
4. **Request transparency from applicants on their use of generative AI, facilitating ways to report it.**
 - Applicants declare if they substantially²¹ used generative AI tool(s) to prepare their application.
 - Applicants provide information on the role of generative AI in the research activities proposed and carried out.
 - Research funding organisations promote an atmosphere where researchers can disclose the use of generative AI transparently and without concerns of being disadvantaged.
 5. **Monitor and get actively involved in the fast-evolving generative AI landscape.**
 - Research funding organisations promote and fund training and educational programmes for an ethical and responsible use of AI in scientific research.
 6. **When using generative AI as part of the internal processes, take into account the risks related to hidden prompts²² in content received from third parties, in particular in applications and proposals.**
 - Research funding organisations remain aware that content may contain hidden prompts or invisible instructions intended to influence generative AI tools.
 - Where relevant, they set clear rules prohibiting such manipulation and introduce appropriate safeguards, in particular in the IT systems involved in the processing of that information.
 7. **Review and evaluate regularly funding instruments and workflows in relation to current AI capabilities and uses.**
 - Research funding organisations are encouraged to review and (re)design their funding instruments to be efficient and resilient to the use of AI, keeping scientific quality and merit at the core of the funding processes.
 - Experimentation to increase the efficiency of research funding organisations' activities is welcome²³. Experimenting with AI demands high levels of transparency with applicants and reviewers, reporting on conclusions and lessons learnt while protecting the rights of applicants (e.g. fair treatment, confidentiality, appealing procedures etc.).

²¹ As mentioned in the recommendations for researchers, using generative AI as a basic support tool for authors is not a substantial use. However, interpreting data analysis, carrying out a literature review, identifying research gaps, formulating research aims, developing hypotheses, etc. could have a substantial impact.

²² A hidden prompt is an instruction secretly embedded in a document or input for an AI system, without being visible to the human reader. Its purpose is to influence how the AI responds, for example by steering it toward a more favourable assessment. In practice, this can be done through invisible text, formatting tricks, or metadata that a person may not notice but an AI tool may still process.

²³ Funders can inspire from CoARA's Experiments in Assessment Idea Catalogue at <https://experiments.coara.org/>.

5. ONGOING DEVELOPMENTS OF GENERATIVE AI IN RESEARCH

AI is developing rapidly. Not all questions about its responsible use in research can be answered yet. Gathering evidence, exchanging practices and building consensus are essential. This section highlights four important developments that may require the Living Guidelines to evolve as our understanding and experience grow.

Agentic AI

Following the OECD's working definition²⁴, an agentic AI system is characterised by: (1) goal complexity, which is the ability to pursue goals that would be challenging for a human to achieve; (2) environmental complexity, which is the ability to operate across domains, time horizons, and external systems; (3) adaptability, the ability to react to unexpected circumstances; and (4) independent execution, which is the ability to reliably act without continuous human intervention. Agentic AI is not a replacement for generative AI: it typically uses one or more generative AI models as its reasoning core (the so-called "harness = model + infrastructure" architecture).

The global climate in which agentic AI is considered as a new enabler in research (fostering productivity and more significant breakthroughs) could potentially erode trust in the results and the credibility of scientific institutions. Agentic AI, by its "autonomous" nature, may obfuscate accountability, and one fundamental principle is that researchers remain the ones ultimately responsible for the quality of the research and the validity of the results. Researchers may inadvertently relinquish or lose control of their work by delegating too many research tasks into agents.

AI agents must be deployed strictly as advanced research tools, not independent scientists. Hypothesis generated, data processed, and conclusions drawn by an AI agent requires validation by human researchers. Ultimately, all ethical responsibility and legal accountability rest solely on the user.

Disclosure of AI use

Disclosure of the use of AI in research practices and publications has been heavily debated. Policies and practices differ across disciplines, regions, publishers and publication systems. Uncertainty about what and how to disclose AI use, administrative reporting burden and fear of negative judgement affect the honest reporting of AI use in science, leading to hesitance and underreporting. In this context, important questions to address are:

- The disclosure threshold: the use of AI is increasingly integrating into the research ecosystem. Which uses of AI warrant disclosure?
- The appropriate level of detail. How to ensure meaningful disclosure without creating excessive reporting burden?
- What is the most appropriate structure, format and placement of scientifically informative AI disclosure in publications?

²⁴ OECD (2026), "The agentic AI landscape and its conceptual foundations", OECD Artificial Intelligence Papers, No. 56, OECD Publishing, Paris, <https://doi.org/10.1787/396cf758-en>.

- How should disclosure of AI use differ along output types (datasets, software, funding proposals, review reports etc.)?

Under the auspices of the International Science Council, COPE, STM and the Global Young Academy, the World Conference on Research Integrity 2026 in Vancouver is developing a global reporting standard for AI disclosure in research. The result of this focus track will inform future versions of these ERA Living Guidelines.

Use of AI in Peer Review

AI is already being used across research, editorial and review workflows, and the central question for peer review is no longer whether it will be used, but how it can be governed responsibly. Peer review remains a fundamentally human process: it validates research through expert scrutiny, strengthens manuscripts through constructive critique, and relies on disciplinary judgement to assess originality, significance, ethics and the strength of evidence. AI may support that process, but it must not replace the accountability, discernment and professional responsibility of reviewers and editors.

These guidelines therefore recommends that generative AI can be used in peer review only in a limited, transparent and responsible manner. Any use should comply with applicable journal, publisher, funder and institutional policies; be disclosed where required; and remain subject to full human responsibility for all assessments and decisions. Confidentiality is non-negotiable: unpublished materials should not be uploaded to AI systems that cannot guarantee high levels of confidentiality, data protection and non-use for training. Secure, trusted, publisher-sanctioned or institutional tools are likely to be necessary for wider adoption and trust.

Before deploying any AI tool, editors and reviewers should first assess whether it is the most suitable and reliable means of performing the specific task, taking into account the nature of the material, the disciplinary context, the need for verifiable provenance, and whether the tool can be used within the applicable confidentiality and accountability guardrails.

On the basis of current experience and observations, the most appropriate uses of AI are supportive and bounded whereas AI should not be used to make or substitute for core peer-review judgements, including assessing novelty, originality, significance, whether evidence supports a research claim, or whether a manuscript should be accepted or rejected.

Use of AI in evaluation of proposals

Generative AI can generate highly polished, optimised and relatively standardised proposal content. This makes proposal preparation much easier and faster, therefore increasing application volumes and overloading funding organisations and evaluators. The optimisation of the application process, for example in the preparation of the non-technical parts (dissemination, intellectual property, etc.) also affects how proposal quality is demonstrated and assessed. It is unclear how the scientific merit of applications improves substantially as, for example, the originality, ambition or scientific value of a proposal remains mostly a responsibility of scientists. The risks of overreliance on generative AI in some parts of the proposal's evaluation processes can impede the entire scientific ecosystem.

These developments underline the continued importance of responsible human judgement: peer review and research assessment remain essential for ensuring scientific quality and trust and should be guided by independence, expertise, fairness, rigour, transparency, confidentiality and accountability.

The fast increase in capabilities of AI systems makes it likely that targeted and minor fixes to the evaluation process will remain one or two steps behind the challenges. The scientific evaluation process therefore probably needs to be redesigned to be more AI-resilient, focusing on aspects that optimise scientific outcomes. At the same time AI could play a role in the evaluation process aiming to increase efficiency and optimise resources in the funding organisations. Experimentation is welcome and needed, while ensuring proper transparency and safeguards.

6. BACKGROUND AND POLICY CONTEXT

ABOUT THESE LIVING GUIDELINES

The widespread and rapid uptake of generative AI has drawn a lot of attention and led to wide policy and institutional responses. The EU is taking the global lead with the AI Act²⁵, and other international governance efforts are taking place. The use of generative AI introduces risks that researchers and organisations need to recognise and manage.

- **Risks to the trustworthiness of research:** AI systems can produce inaccurate or fabricated information, citations or summaries as well as biased or misleading outputs. Uncritical reliance on such outputs can affect the reliability and validity of research.
- **Risks to traceability, accountability and human control:** AI contributions may be difficult to identify or reconstruct, and different AI systems may produce different outputs from the same inputs. This makes it harder to validate or reproduce results, understand how research was conducted or establish responsibility for decisions. Increasingly capable systems can encourage uncritical reliance or inappropriate delegation of scientific judgement ('automation bias').
- **Risks to people, data and privacy:** Personal data, confidential information, unpublished manuscripts, funding proposals or intellectual property may be exposed or processed inappropriately.
- **Systemic risks related to security, openness and autonomy:** Dependence on proprietary or non-open AI systems may create costs, limit access or involve undesirable use of input data. The transfer of critical technology and intellectual property, and the concentration of AI capabilities and infrastructure outside the EU, may create vulnerabilities for the openness, security and strategic autonomy of Europe's research systems.

These guidelines focus on one particular type of AI used in research processes: generative AI along with its evolution towards agentic AI. The guidelines aim to mitigate the aforementioned risks and complement and build on the EU AI policy framework and other policy activities in relation to AI in science, in particular the European AI in Science Strategy, launched in October 2025²⁶.

The development of these guidelines is an important step to prevent misuse and ensure that the use of generative AI has a positive role in research practices. One of the goals of these guidelines is that the scientific community uses this technology in a responsible manner. They are developed with and for the scientific community. All members of the scientific community should actively engage in discussions about the responsible and effective deployment of AI applications, promoting awareness and cultivating a responsible use of AI as part of a research culture based on shared values. Rules and recommendations must go hand in hand with a broad engagement of those involved in public and private research, both organisations

²⁵ <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>

²⁶ https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/european-ai-science-strategy_en

and individuals, to develop a culture of using generative AI in research appropriately and effectively.

Key principles

These guidelines and their underlying principles are based on relevant pre-existing frameworks:

- the European Code of Conduct for Research Integrity²⁷;
- the work and guidelines on trustworthy AI, developed by the High-Level Expert Group on AI²⁸;
- other frameworks, guidance or work developed by the community that we consider relevant for the guidelines.

Why living guidelines?

AI technologies, research practices, evidence and the surrounding policy landscape are in constant development. Some questions concerning the responsible use of AI therefore remain under discussion (see the previous chapter). The ERA Forum will therefore review and update these guidelines as knowledge, practice and consensus develop.

- The first version was published on 20 March 2024.
- This is the fourth version, published on 15 October 2026.
- The next revision is foreseen in the third quarter of 2027.

Researchers, organisations and other stakeholders are invited to contribute experience and evidence that can inform future updates. We welcome your feedback at https://ec.europa.eu/eusurvey/runner/feedback_GenAIResearch.

RESEARCH INTEGRITY

The [European Code of Conduct for Research Integrity](#) by All European Academies (ALLEA) lays down a set of principles to produce sound research, including ethical aspects. These principles include:

- **reliability** in ensuring the quality of research, reflected in the design, methodology, analysis and use of resources;
- **honesty** in developing, undertaking, reviewing, reporting, and communicating research in a transparent, fair, full, and unbiased way;
- **respect** for colleagues, research participants, research subjects, society, ecosystems, cultural heritage and the environment;
- **accountability** for the research from idea to publication, for its management and organisation, for training, supervision, and mentoring, and for its wider societal impacts.

The ALLEA principles in the European [Code of Conduct for Research Integrity](#) serve as guidance for ethical and responsible research. The same value system applies when using generative AI in research. The ALLEA principles has been used as a basis for these

²⁷ [European Code of Conduct for Research Integrity Revised Edition 2023](#)

²⁸ <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

guidelines and be enriched with parts from the other frameworks described in the previous sections.

TRUSTWORTHY AI

The [*Ethics Guidelines for Trustworthy AI*](#) by the EU [High-Level Expert Group on AI](#) focuses on developing responsible AI systems (rather than using them). This work on trustworthy AI contains a set of ethical principles that were translated into operational requirements for developing AI systems. The requirements has been a useful inspiration for drawing up guidelines on the use of generative AI.

The four ethical principles for AI systems are:

1. respect for human autonomy;
2. prevention of harm;
3. fairness;
4. explicability.

These ethical principles were used to develop these seven operational key requirements:

1. human agency and oversight²⁹;
2. technical robustness and safety;
3. privacy and data governance;
4. transparency;
5. diversity, non-discrimination and fairness;
6. environmental and societal well-being;
7. accountability.

²⁹ Human agency and oversight include human-in-the-loop (HITL), human-on-the-loop (HOTL) and human-in-command (HIC) approaches. HITL involves human intervention in every decision cycle of the system, which in many cases is neither possible nor desirable. HOTL involves human intervention during the design cycle of the system and monitoring how the system operates. HIC involves overseeing the overall activity of the AI system (including its broader economic, societal, legal and ethical impact) and the ability to decide when and how to use the system in any particular situation (Ethics Guidelines for Trustworthy AI).

REFERENCES

ARTICLES AND PAPERS

Birhane, A., Kasirzadeh, A., Leslie, D. & Wachter, S., 'Science in the Age of Large Language Models', *Nature Reviews Physics*, Vol. 5, 2023, pp. 277-280.

<https://doi.org/10.1038/s42254-023-00581-4>

Conroy, G., 'Scientific Sleuths Spot Dishonest ChatGPT Use in Papers', *Nature*, 2023.

<https://doi.org/10.1038/d41586-023-02477-w>

Conroy, G., 'How ChatGPT and Other AI Tools Could Disrupt Scientific Publishing', *Nature*, Vol. 622, 2023, pp. 234-236. <https://doi.org/10.1038/d41586-023-03144-w>

Ganjavi, C., Eppler, M., Pekcan, A., Biedermann, B., Abreu, A., Collins, G., Gill, I. Cacciamani, G., 'Bibliometric Analysis of Publisher and Journal Instructions to Authors on Generative-AI in Academic and Scientific Publishing', *arXiv*, 2023. <https://doi.org/10.48550/arXiv.2307.11918>

Lenharo, M., 'ChatGPT turns two: how the AI chatbot has changed scientists' lives', *Nature*, Vol. 636, pp. 281-282 (2024). <https://doi.org/10.1038/d41586-024-03940-y>

Li, Z., Liang, C., Peng, J., Yin, M., 'How Does the Disclosure of AI Assistance Affect the Perceptions of Writing?' *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 4849–4868 (2024). [10.18653/v1/2024.emnlp-main.279](https://arxiv.org/abs/2408.11918)

Lorenz, P., Perset, K. & Berryhill, J., 'Initial Policy Considerations for Generative Artificial Intelligence', *OECD Artificial Intelligence Papers*, No. 1, OECD Publishing, Paris, 2023. <https://doi.org/10.1787/fae2d1e6-en>

Lund, B.D., Wang, T., Mannuru, N.R., Nie, B., Shimray, S. & Wang, Z., 'ChatGPT and a New Academic Reality: Artificial Intelligence-Written Research Papers and the Ethics of the Large Language Models in Scholarly Publishing', *Journal of the Association for Information Science and Technology*, Vol. 74, No. 5, 2023, pp. 570-581. <http://dx.doi.org/10.1002/asi.24750>

Nogueira, L. A., Moltubakk, S. T., Fagervik, A., & Langfeldt, I. B. 'Cutting through the noise: Assessing tools that employ artificial intelligence.', *IFLA Journal*, 0(0), 2025 <https://doi.org/10.1177/03400352241304121>

Parrilla, J.M., 'ChatGPT Use Shows that the Grant-Application System is Broken', *Nature*, 2023. <https://doi.org/10.1038/d41586-023-03238-5>

International Association of Scientific, Technical and Medical Publishers (STM), 'AI Ethics in Scholarly Communication: STM Best Practice Principles for Ethical, Trustworthy and Human-centric AI', 2021

<https://coilink.org/20.500.12592/hqbzq79>

Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askill, A., Bowman, S. R., ... & Perez, E. (2023). 'Towards understanding sycophancy in language models.' *arXiv*

<https://doi.org/10.48550/arXiv.2310.13548>

UKRI, *Summary: the roles of research organisations, publishers and funders in upholding responsible Artificial Intelligence use* <https://ukrio.org/wp-content/uploads/20260805-AI-and-peer-review-roundtable-output-PDF.pdf>

Van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL., 'ChatGPT: Five Priorities for Research', *Nature*, Vol. 614, pp. 224-226 (2023) <https://doi.org/10.1038/d41586-023-00288-7>

Van Noorden R. & Perkel J., 'AI and Science: What 1,600 Researchers Think', *Nature*, Vol. 621, pp. 672-675 (2023) <https://doi.org/10.1038/d41586-023-02980-0>

United Nations Environment Programme, 'Artificial Intelligence (AI) end-to-end: The Environmental Impact of the Full AI Lifecycle Needs to be Comprehensively Assessed - Issue Note' (2024) <https://wedocs.unep.org/handle/20.500.11822/46288>

Zou, J., 'ChatGPT is transforming peer review — how can we use it responsibly?', *Nature* Vol. 635, p.10 (2024) <https://doi.org/10.1038/d41586-024-03588-8>

GUIDELINES ANALYSED

Aalto University:

<https://www.aalto.fi/en/services/responsible-use-of-artificial-intelligence-in-the-research-process>

All European Academies (ALLEA): <https://allea.org/wp-content/uploads/2023/06/European-Code-of-Conduct-Revised-Edition-2023.pdf> (DOI: 10.26356/ECOC)

Arizona State University:

<https://researchintegrity.asu.edu/export-controls-and-security/artificial-intelligence>

Association for Computing Machinery (ACM):

<https://www.acm.org/binaries/content/assets/public-policy/ustpc-approved-generative-ai-principles>

Australian Government Tertiary Education Quality and Standards Agency (TEQSA):

<https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/artificial-intelligence>

Bockting CL, van Dis EAM, van Rooij R, Zuidema W, Bollen J. Living guidelines for generative AI - why scientists must oversee its use. *Nature*. 2023 Oct;622(7984):693-696.

<https://doi.org/10.1038/d41586-023-03266-1>

Committee on Publication Ethics (COPE):

<https://publicationethics.org/cope-position-statements/ai-author>

Deakin University: <https://deakin.libguides.com/generative-AI/ethics-evaluation>

Deutsche Forschungsgemeinschaft (DFG):

https://www.dfg.de/download/pdf/dfg_im_profil/geschaeftsstelle/publikationen/stellungnahmen_papiere/2023/230921_statement_executive_committee_ki_ai.pdf

Elsevier: <https://www.elsevier.com/about/policies/publishing-ethics>

European Research Council, The use of AI in grant proposal evaluation,

<https://erc.europa.eu/system/files/2026-03/Use-AI-grant-proposal-evaluation.pdf>

ETH Zurich: <https://www.research-collection.ethz.ch/handle/20.500.11850/672207>

European Commission, Guidance note — Instructions for the use of AI tools by experts
https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/experts/guidance-note-instructions-for-use-of-AI-tools-by-experts_en.pdf

Harvard University:
<https://provost.harvard.edu/guidelines-using-chatgpt-and-other-generative-ai-tools-harvard>

International Association of Scientific, Technical and Medical Publishers (STM):
[Global Principles for AI](#)

International Association of Scientific, Technical and Medical Publishers (STM):
[Ethical and practical guidelines for the use of generative AI in the publishing process](#)

International Society for Computational Biology (ISCB):
<https://www.iscb.org/iscb-policy-statements/iscb-policy-for-acceptable-use-of-large-language-models>

KU Leuven: <https://research.kuleuven.be/en/integrity-ethics/integrity/practices/genai/genAI>

Maastricht University: <https://www.maastrichtuniversity.nl/news/guidelines-use-generative-ai>

Science Europe:
<https://www.scienceeurope.org/our-priorities/research-culture/research-values-framework>

United Nations Educational, Scientific and Cultural Organization (UNESCO):
<https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>

University of Ljubljana: <https://www.uni-lj.si/mma/uien/2023092013475160/?m=1695210471>

University of Toronto:
<https://www.sgs.utoronto.ca/about/guidance-on-the-use-of-generative-artificial-intelligence/>

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: **00 800 6 7 8 9 10 11** (certain operators may charge for these calls),
- at the following standard number: **+32 22999696**,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (europa.eu).

EU Publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

EU open data

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

These guidelines are the result of a joint effort between the European Commission and the countries and stakeholders represented in the ERA Forum. They provide simple and actionable recommendations to key actors in the research community on the use of generative AI. These recommendations will bring clarity and reassurance to those using generative AI in their research activities.

These are ‘living’ guidelines, they will be updated regularly to keep up with the very fast technological development in this area.

Research and Innovation policy



**Publications Office
of the European Union**

Law | Data | Publications