Summary note

Uses of generative Artificial Intelligence tools in research
- Points for attention and best practice -

Céline Blitz Frayret, UMR Eco&Sols

Estelle Jaligot, Ethics and research integrity office

Colline Orsini, Legal affairs and compliance office

Table des matières

Disclaimer: the translation of this text from French into English was improved using the AI-based tool DeepL.

I Introduction

I.1 Subject of the note

This note is an effort to summarize the main issues raised by new generative Artificial Intelligence (GenAI) tools and their applications in scientific research activities, with the aim of drawing attention to the potential risks (whether ethical, legal, scientific, social or environmental) associated with their use.

Our recommendations should be understood as countermeasures to reduce or eliminate the risks presented by current GenAIs. They are not based on an exhaustive and strict list of "permitted" or "prohibited" uses. Indeed, such a list would be obsolete as soon as drawn up and could never cover all cases of use of all GenAI models. In contrast to such an approach, **we provide users with guidelines and avenues for reflection, enabling them to use AIg in a responsible and enlightened way, or to choose not to do so in full knowledge of the facts.**

Although the examples used to illustrate our points are drawn primarily from scientific applications of GenAI, some of them may also concern administrative and support staff in research establishments. For this reason, we have opted to use language that is as accessible to non-specialists as possible.

This document is completed by a list of the resources we have used: they deal with AI in general or GenIA in particular, and they either have a broad or a thematic scope (**Appendix 1**).

This note will evolve as new uses of GenAIs are assessed and their limitations and biases analyzed, opening the door to improvements in their design and performance. Any contribution to enrich it is therefore welcome.

I.2 GenAIs: what are they?

Recently, new AI tools based on large language models (LLMs) have been made available to the general public to considerable publicity: ChatGPT (offered by OpenAI) is the best-known of these, but there are also Bard and Gemini (Google), Llama (Meta), Claude (Anthropic), and many others.

What all these tools share is the fact that they are conversational agents (chatbots), reacting to a prompt from the user to produce a credible, if not correct, response. They operate on the basis of an algorithm which performance has been progressively refined and validated by human annotators on large amounts of training data. The answer provided is a synthesis of pre-existing information that has been either supplied by the user as a complement to the prompt or drawn by the GenAI from the resources to which it has access.

In the present note, the issues we have chosen to highlight are mainly related to the use of textual GenAIs, due to their very broad application potential in a search context. GenAI tools that are dedicated to image (e.g. Dall-E, Midjourney, Firefly) or sound generation (AudioCraft, AIVA, SOUNDRAW, etc.), as well as general-purpose AIs such as Copilot (Microsoft) share some of these problems.

The AI Act: an assessment of its implications for research is needed

The rapid growth of AI has prompted Europe to produce a long-awaited piece of legislation: the AI Act, which will come into force in June 2026. This regulation adopts a **risk-based approach**: in practical terms, the more potential risks AI carries, the greater the obligations (most of which weighing on the supplier) will be. It also clarifies **the roles and responsibilities of the various players involved in AI systems**, stresses the importance of system transparency and prohibits certain practices.

As far as **research** is concerned, the regulation includes an **exclusion** designed to preserve the capacity for innovation: it stipulates that **the obligations laid down in the AI Act will not apply "*to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development*"**. However, an in-depth study, based on the examination of a wide range of use cases, will be necessary to determine the scope of application of this exception. Indeed, while it seems clear that it is aimed at AI models developed and applied in the context of a research project, the status of AIs designed to support research is less clear. Conversely, AI systems used to conduct research and development but designed for other purposes will be subject to the AI Act[1]. This applies, for example, to the use of any commercial GenAI model in a research project.

The regulation also recalls that recognized ethical and professional standards in scientific research and EU law remain applicable, regardless of whether the AI model is covered by the AI Act or not.

[1] This is specified in the 25th recital of the AI Act.

II GenAI: known issues and implications for research

II.1 Production or propagation of erroneous or incomplete data

GenAIs produce answers based on statistical correlations that are based on the frequency of association between strings of characters, which may have no bearing on their meaning or scientific validity. The tool does not have the capacity to critically analyze the information processed or produced, and is therefore unable to interpret it, prioritize it or assign a totally reliable level of confidence to it. **GenAI tools are therefore liable to propagate or produce false information**.

Furthermore, the datasets used to train the GenAI may also be of poor quality due to their incomplete or erroneous nature (whether these errors are of an unintentional nature or the result of deliberate "data poisoning"), increasing the likelihood of producing inadequate outputs.

In the context of research, GenAI can produce outputs that are unrelated to real-world information by rearranging or reformulating unrelated pre-existing elements (so-called **"hallucinations"**). The ability of GenAI to produce combinations of true and false assertions on a given subject is acknowledged by developers[2]. Thus, GenAIs will almost always produce an answer, however inaccurate or inadequate it may be in relation to the initial prompt: indeed, **few GenAIs mention their inability to answer or their uncertainty**.

While they are able to accelerate the writing of lines of computer code, GenAIs are likely to produce errors leading to **data loss** or even **security breaches**.
Since GenAIs have to "learn" from available data before they can provide an answer, they are **"blind" to the most recent information**: thus, a state of the art that is produced by a GenAI tool will have to be updated manually by the user.

II.2 Production of variable and non-reproducible responses

Because of the way they are developed and operate, GenAIs tend to produce variable responses to the same query, which raises the question of the reproducibility of the results obtained with these tools. The main factors of variation identified are as follows:

- The **modalities of interaction** with the GenAI tool (use of one or more prompts, formulation and order of their use) are likely to induce significant variations in the outputs.

- **GenAIs performance evolves continuously** as a result of i) updates from the developers and ii) successive user requests, which contribute to the training process.

- **The language model underlying a GenAI tool is developed independently for each language** using a corpus of resources, some of which may be specific to that language. As a result, the outputs generated by the GenAI tool may differ (including in terms of quality) from one language to another due to variations in the size of the corpus of data available and the number of queries made in each language. As far as research is

[2] Hence the following disclaimer under ChatGPT's window: "ChatGPT may produce inaccurate information about people, places, or facts".

concerned, the quality of the answers obtained will therefore be much higher in English, the language of international exchanges, than in a less widely used language.

II.3 Representational bias

Since they are based on pattern recognition within a text, GenAIs tend to preferentially reproduce word and sentence sequences according to their strongest representation within the data used for training and/or that provided by the user. They will therefore tend to provide answers based on elements (facts, opinions) that are in the majority within these data, and to avoid minority elements. This mode of operation may lead to **stereotyped responses** that miss important nuances or generate misunderstandings.

This problem can be all the more critical as the development of the algorithm and the training datasets may themselves be biased. GenAIs will then provide answers that **reproduce** or even **amplify these biases**. A decision made on the basis of GenAI-generated outputs can therefore lead to **inequitable situations, or even discrimination**[3].

II.4 Issues related to source attribution and citation

GenAIs do not create any intrinsically new material; their outputs are produced through the rearrangement of pre-existing elements. Insofar as the operation of GenAIs may rely on the (sometimes unauthorized) exploitation of **copyrighted material** (whether in training or user-supplied data), the outputs constitute an **infringement** of these rights[4]. In the case of the production of scientific writings or illustrations, the use of a GenAI tool can thus be considered as **plagiarism** or even **counterfeiting**.

Similarly, the fact that most GenAIs do not mention the authorship of the original works in their outputs breaches the principles of **copyright** (moral rights, economic rights). By extension, the user is at fault when using their outputs. The regulatory texts addressing this subject place great emphasis on the risks of copyright infringement posed by GenAIs. For example, the Hiroshima G7 guidelines recommend the implementation of appropriate measures to seize and protect intellectual property. The AI Act also makes respecting and protecting these rights a necessity.

In the context of research, the use of most GenAIs is problematic due to their **inability to cite their sources**, or their tendency to **cite them inaccurately or out of context**. The production of "hallucinated" bibliographic references has been reported in many instances.

II.5 Issues related to the reuse of user-supplied data

In most cases (and especially for online models that are available free of charge), user-supplied data is then integrated into the operation of the GenAI tool, and is therefore likely to be accessible to any user through future outputs. Wherever the data supplied to the GenAI tool is

[3] The lack of diversity and inclusiveness within the digital sector may be an aggravating factor of this phenomenon.
[4] This exploitation can occur without the authors' knowledge, even when it is strictly speaking legal. Recently, for example, two partnership agreements involving an academic publisher and a private group gave the latter the possibility of using the contents of the former's publications catalog to train its AI models (source: The Chronicles of Higher Education, 29 July 2024).

not publicly accessible (e.g. internal documents), the use of such a model may therefore lead to **data leakage**.

In a research context, the use of this type of open GenAI tool in the drafting or formatting of confidential documents therefore constitutes a **breach of the duty of confidentiality**. If this obligation of confidentiality has been included in a contract, such disclosure runs counter to the **contractual commitments**, which may engage the responsibility of the user or that of the institution.

By the same logic, the use of such GenAIs on documents containing personal data breaches the **obligation to ensure the protection of personal data and/or privacy**, in accordance with the General Data Protection Regulation (GDPR)[5].

More broadly, this leakage of data raises the question of their **sovereignty**, since the user will have no visibility over the uses that might be made of them by the publisher and/or host of the GenAI tool. This is all the more true as these uses are potentially subject to **legal rules that differ from those applicable in France**. For example, ChatGPT is a tool subject to American laws, which are based on a very broad definition of the concept of "sovereignty"[6]. The AI Act insists on the need to respect the protection of confidential commercial information and trade secrets.

II.6 Lack of transparency for users

Because of their complexity and the sheer volume of data they process, GenAIs are extremely opaque when it comes to the way they work: they are often likened to **"black boxes"**. As a result, users generally have neither the means to prevent the above-mentioned problems, nor to exercise real quality control. As a result, **it may be difficult (if not impossible) to explain how the GenAI tool arrived at a result or to interpret it**. This is especially true if users are unfamiliar with the use of AI-based tools. This phenomenon is further accentuated in the case of GenAI tools developed by private companies who protect their innovations from the competition through secrecy.

In the context of research, this lack of transparency deprives the user of the possibility of fully guaranteeing the integrity of the data produced and the rigor of the process. As this obligation is one of the **basic responsibilities of any author of a scientific publication**, a growing number of scientific journals and publishers have introduced guidelines that strictly regulate the use of GenAIs in publications. While these guidelines are still very heterogeneous[7], they generally share the common features of i) emphasizing the **author's complete responsibility for the results produced by any GenAI tool** (which cannot be credited as author nor cited as source, since it lacks the capacity to assume such responsibility) and ii) requiring **full disclosure** of these uses.

[5] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016.
[6] The Cloud Act allows US authorities with a warrant to access the data of individuals and companies located outside the USA, provided that the entity hosting the data is based in the USA or is of US nationality.
[7] The CANGARU initiative aims to elaborate guidelines for GenAI reporting and use in publications by integrating the perspectives of various stakeholders and different scientific fields (see reference in **Appendix 1**).

II.7 Lack of openness and traceability of operations and outputs

In line with the **principles of open science**, access to the processing chain of GenAI-generated outputs (including AIg source codes, data, etc.) should be allowed to enable the reproduction of the result in the event of an audit or challenge. However, in most cases the GenAI model does not cite its sources, making verification difficult. Beyond the verification of the veracity of the answers produced, this opacity of the tool's design and operation prevents the user from satisfying the **FAIR principles** of good research data management[8].

As GenAIs are designed to emulate human creations, **the detection of their outputs** (by other AIs or by humans) **is as yet unreliable**. This inability to reliably distinguish legitimate information from that produced by an AI therefore places the onus on the user to declare his use of the GenAI tool, so that neither the origin of their data nor the rigor of their work may be questioned. In the context of a research activity, **the undisclosed use of an AI constitutes a breach of scientific integrity**[9].

Ethical concerns are complemented by a growing number of legal texts (practical information sheets from the Commission Nationale de l'Informatique et des Libertés, the AI Act, etc.), which underline the importance of transparency in the tools developed by suppliers.

II.8 Facilitation of voluntary or involuntary research misconducts and other unacceptable practices

As mentioned above, the inappropriate use of GenAIs may facilitate **plagiarism**. The ability of these tools to mimic existing data can also be misused to generate fictitious research data (based on no experimentation, observation or simulation: **fabrication**) or to manipulate real data in order to alter its interpretation (**falsification**). All three are instances of **research misconduct**[10]. In recent years, the **large-scale exploitation of GenAIs by unscrupulous mercantile players (papermills)** has led to the mass publication of scientific articles based on fabricated, falsified or plagiarized data. Contributing, directly or indirectly, to the activities of a papermill (through the purchase of an author position on a fraudulent article or by allowing its publication as an editor or reviewer) is a **breach of scientific integrity**[11].

In the long term, the growing proportion of these fraudulent articles is likely to cast doubt on the validity of legitimate research results, and thus **"pollute" the body of knowledge** on which subsequent research is based. In the longer term, such scientific "fake news" may have catastrophic **societal impacts** (implementation of erroneous public policies, release of poorly performing or even harmful products, waste of resources, etc.) and **undermine society's trust in science and scientists**.

---

[8] Findable, Accessible, Interoperable, Reusable: see Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18.

[9] **European Code of Conduct for research integrity**, section **Research Misconduct and other unacceptable practices:** "*Hiding the use of AI or automated tools in the creation of content or drafting of publications*" (see reference in **Appendix 1**).

[10] Source: see note 9.

[11] "*Establishing, supporting, or deliberately using journals, publishers, events, or services that undermine the quality of research ('predatory' journals or conferences and paper mills).*" Source: see note 9.

Aside from these exceptional cases, **less severe integrity breaches** may emerge because of the highly heterogeneous levels of proficiency in the responsible use of GenAIs and the variable accessibility of the resources needed to improve practices. There is therefore a strong need to simultaneously **build capacity** in a way that caters to the diversity of users, uses and contexts, and to **homogenize practices**.

III GenAIs: impacts beyond research

The development, training and operation of GenAIs require the manipulation of comparatively **larger volumes of data** than for other AI models. Moreover, their **great versatility** make them suitable for use by a wide range of people for a wide range of purposes. For this reason, some of the problems experienced with other AIs are amplified in the context of GenAIs.

Thus, AIs and GenAIs especially:

- Have a significant **environmental footprint** in terms of energy and water consumption, as well as greenhouse gas emissions - most of which are generated by data centers. For example, a query submitted to a GenAI tool has an environmental impact that is 5- to 10-folds higher than that of a query processed by an online search engine.

- Are potentially driving profound **transformations in the labor market**, as well as in the conditions under which many professions operate. Worldwide, it is estimated that 40% of professions are highly exposed to AI, whether it will replace or support their activities. While productivity gains and higher incomes are anticipated for certain professions, the implementation of support measures for categories of workers for whom the AI adoption is more difficult (due to their age, level of qualification, the specificities of their profession, etc.) will be essential in order to avoid accentuating economic and social inequalities.

- Exploit and potentially accentuate **structural inequalities between countries from the global North and South**:
  - through their massive use of low-paid workers from the South who are sometimes employed under unethical conditions;
  - because of the significant differences in AI adoption readiness indicators[12] between countries with different levels of economic development;
  - as a result of the sector's heavy dependence on technologies and infrastructures owned by large private groups (notably multinationals such as the GAFAM[13]) based in high-income countries.

- Entail very high **operating costs**[14], making their use unaffordable and, in the long term, economically unsustainable.

This raises questions about the **economic model** that GenAIs implicitly promote, and about the ability of governments (particularly those in the global South) to guarantee their autonomy and sovereignty in the matter.
Moreover, managing the specific modalities of interaction with GenAIs and understanding their potential and limitations may require specific support and training. Differential access not only to GenAI tools but also to this support (depending on socio-cultural and economic contexts, etc., or according to scientific professions or disciplines in the case of research) may therefore accentuate or give rise to **inequalities**.

[12] IMF's AI Preparedness Index (AIPI) summarizes indicators related to digital infrastructure, human capital and labor market policies, innovation and economic integration, and regulation and ethics (see reference in **Appendix 1**).
[13] Google/Alphabet, Amazon, Facebook/Meta, Apple, Microsoft.
[14] A query submitted to ChatGPT has an estimated financial cost of 500 US dollars.

IV Good practice for the use of GenAIs in research

GenAIs are powerful tools for rapidly gathering, synthesizing and reformulating large quantities of information. From this perspective, they can be useful for **increasing the efficiency** of certain tasks, especially low-complexity ones. GenAIs may also help to **improve inclusiveness** by providing diverse audiences easier access to scientific information.

The responsible use of GenAI tools involves:

- **Always checking whether using a GenAI tool is allowed and how it should be used.** Consult the guidelines for the intended use (e.g. instructions to authors, recommendations to reviewers, funder's policy) or take advice from reference persons in the field (group or project leader, journal editor, review committee or jury chair, etc.).

- **Always being fully transparent when using an AI tool (generative or otherwise)**: minimum information to be disclosed should include the model, version, source and rationale. Whenever possible, make explicit and accessible the strategies used in the interation with the GenAI model (formulation of the prompt or succession of prompts) as well as the knowledge used (nature and modality of access to source data), while respecting the legal and ethical framework of the latter. If this option is available, choose a GenAI tool which source code is open.

- **Never using a GenAI tool to produce all or part of any material presented as original work for which either the author or the employing or hosting institution may be held accountable.** This applies to the production of research datasets, scientific publications, expert or assessment reports, policy briefs, funding applications, dissertations or theses, etc.

- **Where permitted, using GenAI tools as writing assistants** as long as this is limited to occasional support for reformulation, correction of syntax or spelling, or improvement of style. However, it is advisable to keep this use to a strict minimum, for example, by using the tool to edit small portions of text written by the user.

- **Limiting the use of GenAIs to the production of outputs for which the user has the necessary knowledge and skills to carry out a quality and validity check, for which they will be held responsible.** This could involve, for example, producing a summary (based on documents to which they have access), reformulating a translation (between two languages with which they have a reasonable command) or validating a computer program (in a language and for functions with which they are familiar).

- **Never feeding an open GenAI with data that is not intended to be made public.** This includes:
  - **personal data** (e.g.: contact lists, CVs, professional, social or medical records, etc.);
  - **confidential data** (unpublished research data, manuscripts or research proposals submitted for evaluation, minutes of decision-making bodies, contractual documents, internal documents relating to the institution's strategy or operations, etc.);

- **data protected by the intellectual property rights of a third party who has not given their consent** (texts, images or sounds that you do not own, publications whose copyright has been transferred to the publisher).

Wherever possible, it is preferable to use a **"business" version** of the GenAI model rather than the free public version. Indeed, such a version is adapted to professional use and generally offers protection against the risk of data leakage. If a **sovereign GenAI** model (i.e. one that has been developed by either the French government or a national public structure) covers the intended uses, it should be preferred in all cases. **If no "business" or sovereign model can be used, the user should consider remove non-public data (as described above) from their exchanges with the GenAI tool, or refrain from using it.**

- **Whenever a GenAI tool is used to support decision-making, treating the answers provided as one element among others** and comparing them with elements obtained without the tool. Subjecting, in all cases, both the process and the output itself to **human supervision** and **critical analysis** before any decision is made.

- **Contributing**, as far as possible, **to discussions and to the definition of good practices** for the use of GenAIs, so that they are tailored to a wide variety of users, fields of activity, contexts, etc.

- Ensuring that use of GenAIs by **partners, collaborators or subcontractors** complies with best practices.

- As the head of an institution or group, **preventing the emergence or accentuation of inequalities in access to and proficiency in GenAI tools**. Taking into account any differences in levels of knowledge of the benefits and risks of AIg (depending on scientific disciplines, professions, types of stakeholders - research or civil society -, categories of staff, geographical or cultural contexts, etc.). Promoting capacity building for staff while providing specific support for people or categories likely to encounter difficulties in adopting GenAIs (due to their age, gender, education level, etc.).

- As a **stakeholder in scientific publishing** (editing, reviewing), **preventing the undisclosed or abusive use of GenAI tools**, especially in the context of papermills: familiarizing or training oneself to the use of detection techniques and tools, seeking support from experts, raising awareness on these issues.

- In the context of a research concerned about its impact on people, society and the environment, **assessing the relevance of using GenAI tools against the use of non-AI tools**. When their use proves to be sufficiently justified, rationalizing it in relation to the risks identified, and putting in place suitable measures in order to minimize any negative impacts.

- **Resisting the temptation to use GenAIs opportunistically (without rigorous scientific justification) or as a workaround for a shortage of resources or skills.**

V Appendix 1 : Resources relative to the use of GenAIs in research

V.1 Resources of general interest

National level (France)

- *Commission Nationale de l'Informatique et des Libertés (CNIL):*

Les fiches pratiques sur l'IA

LINC - Laboratoire d'innovation numérique de la CNIL: Dossier "IA générative"

- *Comité national pilote d'éthique du numérique (CNPEN):*

Avis n°3: Agents conversationnels: enjeux d'éthique, 15 septembre 2021.

Avis n°7: Systèmes d'intelligence artificielle générative : enjeux d'éthique, 30 juin 2023.

- *Association Data for Good:*

Livre blanc "Les grands défis de l'IA générative". Version 1.0, juillet 2023.

- *Office français de l'intégrité scientifique (OFIS):*

Systèmes d'intelligence artificielle générative : quelques points de vigilance, février 2024.

- *Académie Nationale de Médecine:*

Systèmes d'IA générative en santé : enjeux et perspectives, rapport adopté le 5 mars 2024.

Other countries

- The Royal Society (UK):

Science in the age of AI. Report, May 2024.

- UK Research integrity office (UKRIO):

AI in research (updated January 2024).

- *Pôle Interordres de Montréal & Laboratoire d'éthique du numérique et de l'Intelligence Artificielle (LEN.IA), Québec:*

Former à l'éthique de l'IA en enseignement supérieur: référentiel de compétences; trousse pédagogique

International level

- *United Nations Educational, Scientific and Cultural Organization (UNESCO):*

Recommendation on the Ethics of Artificial Intelligence (23 November 2021).

- *Organization for Economic Co-operation and Development (OECD):*

Recommendation of the Council on Artificial Intelligence (22 May 2019).

OECD.AI policy observatory - Policies, data and analysis for trustworthy artificial intelligence (website).

- *G7*

Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems, 30 October 2023

Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems, 30 October 2023.

- Montreal AI Ethics Institute (MAIEI):

Website.

- *European Parliament and Council :*

AI Act: Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence

- *European Commission (EC):*

Ethics guidelines for trustworthy AI, 2019.
Ethics By Design and Ethics of Use Approaches for Artificial Intelligence, version 1.0, 25 November 2021.
Living guidelines on the responsible use of generative AI in research, version 1.0, March 2024.

- *All European Academies (ALLEA):*

European Code of Conduct for research integrity (2023 revised edition).

V.2 Thematic resources

GenAIs in scientific publications

Research, investigations, reports:

Bhavsar D, Duffy L, Jo H, Lokker C, Haynes RB, Iorio A, Marusic A, Ng JY (2024) Policies on Artificial Intelligence Chatbots Among Academic Publishers: A Cross-Sectional Audit. medRxiv preprint: https://www.medrxiv.org/content/10.1101/2024.06.19.24309148v1
Cacciamani GE, Eppler MB, Ganjavi C, Pekan A, Biedermann B, Collins GS, Gill IS (2023) Development of the ChatGPT, Generative Artificial Intelligence and Natural Large Language Models for Accountable Reporting and Use (CANGARU) Guidelines. arXiv preprint: http://arxiv.org/abs/2307.08974
Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, Gill IS, Cacciamani GE (2024) Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. BMJ 384:e077192. https://doi.org/10.1136/bmj-2023-077192
Kacena MA, Plotkin LI, Fehrenbacher JC (2024) The Use of Artificial Intelligence in Writing Scientific Review Articles. Curr Osteoporos Rep 22:115–121. https://doi.org/10.1007/s11914-023-00852-0
Liang W, Zhang Y, Cao H, Wang B, Ding DY, Yang X, Vodrahalli K, He S, Smith DS, Yin Y, McFarland DA, Zou J (2024) Can Large Language Models Provide Useful Feedback on Research Papers? A Large-Scale Empirical Analysis. NEJM AI 1:AIoa2400196. https://doi.org/10.1056/AIoa2400196
Mugaanyi J, Cai L, Cheng S, Lu C, Huang J (2024) Evaluation of Large Language Model Performance and Reliability for Citations and References in Scholarly Writing: Cross-Disciplinary Study. Journal of Medical Internet Research 26:e52935. https://doi.org/10.2196/52935
Thelwall M (2024) Can ChatGPT evaluate research quality? Journal of Data and Information Science 9:1–21. https://doi.org/10.2478/jdis-2024-0013

Recommendations:

- *International Committee of Medical Journal Editors (ICMJE):*

"Defining the role of authors and contributors" - 2024 update including a point on the use of GenAI tools

- *Committee on Publication Ethics (COPE):*

Authorship and AI tools - COPE position statement, February 2023.

- *International Association of Scientific, Technical, and Medical Publishers (STM):*

Generative AI in scholarly communications: ethical and practical guidelines for the use of generative AI in the publication process, December 2023.

Databases - Retraction Watch:
Papers and peer reviews with evidence of ChatGPT writing.

Press:
How ChatGPT and other AI tools could disrupt scientific publishing (Nature, 10 October 2023).
AI-generated rat genitalia: Swiss publisher of scientific journal under pressure (SWI swissinfo.ch, 13 March 2024).
Quand ChatGPT tient la plume (TheMetaNews, 26 avril 2024).

Anti-papermills platforms:
STM Integrity Hub
UNITED2ACT

Social and environnemental impacts of AIs
Research, investigations, reports:
Li P, Yang J, Islam MA, Ren S (2023) Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. arXiv preprint: http://arxiv.org/abs/2304.03271
Ludec CL, Cornet M (2023) How low-paid workers in Madagascar power French tech's AI ambitions. The Conversation.

Topic:
International Monetary Fund (IMF): "Artificial Intelligence"

Press:
"Ils profitent de notre pauvreté" : derrière le boom des intelligences artificielles génératives, le travail caché des petites mains de l'IA (France Info, 8 avril 2024).
IA : quel est le bilan carbone de ChatGPT ? (Les Numériques, 22 avril 2024).
À Madagascar, les petites mains bien réelles de l'intelligence artificielle (France Info, 29 avril 2024)
Comment l'intelligence artificielle a fait augmenter les émissions de gaz à effet de serre des géants de la tech (France Info, 3 juillet 2024).
IA: la consommation d'eau cachée de ChatGPT (Disclose, newsletter "Planète investigation", 4 juillet 2024).