European Commission

Research and Innovation

The Use of Generative Artificial Intelligence in Research

R&I PAPER SERIES
POLICY BRIEF

The Use of Generative Artificial Intelligence in Research

European Commission
Directorate-General for Research and Innovation
Directorate G — Common Policy Centre
Unit G.1 — Common strategy & Foresight Service
Contact Athina Karvounaraki, Alexis Stevenson
Email athina.karvounaraki@ec.europa.eu
alexis.stevenson@ec.europa.eu
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in June 2025

PDF	ISBN 978-92-68-26702-8	doi: 10.2777/1024414	KI-01-25-084-EN-N

Luxembourg: Publications Office of the European Union, 2025

EUROPEAN COMMISSION

The Use of Generative Artificial Intelligence in Research

Isabelle Labrosse & David Campbell, Science-Metrix
Henrik Karlstrøm & Eric Iversen, NIFU
Lili Wang & Ad Notten, UNU/MERIT, Maastricht University

2025 Directorate-General for Research and Innovation EN

Contents

1. Highlights and recommendation 3

2. Background 4

3. Literature review 4

3.1. The Benefits of GenAI for Research 6

3.2. The Risks and Ethics of GenAI for Research 6

3.2.1. Risk Awareness 6

3.2.2. Ethical Implications 7

3.3. Cross-impact on Education 9

4. Data Analysis 10

4.1. Trends in GenAI usage in Research 10

4.2. Context of GenAI Chatbot Mentions 11

4.2.1. Documents with GenAI chatbot mentions by research area and their use cases 12

4.2.2. Documents with GenAI chatbot mentions by document type and section, as well as their use cases 15

4.3. Traces of GenAI Assistance in Scientific Publishing 18

5. Discussion 20

6. Bibliography 22

7. Annex – Study Methodology 26

7.1. Declaration of generative AI and AI-assisted technologies in the working process 26

7.2. Literature Review 26

7.3. Quantitative Analysis 27

7.3.1. Identification of publications with mentions of GenAI chatbots 27

7.3.2. Context of use of GenAI 28

7.3.3. Occurrence of AI-favoured words in publications 28

1. Highlights and recommendation

Despite early efforts by a range of relevant stakeholders to raise awareness of the risks and ethical implications of generative AI (GenAI) in research, these efforts remain heterogeneous across the entire research landscape and have not reached all spheres of academia. Accordingly, there is general agreement that the academic community and policy makers, both at the national and supranational levels, must develop and adopt **common definitions, ethical frameworks, policies and guidelines** on the areas in which it can be responsibly used and regarding the integration of these technologies into scientific practice and reporting to prevent malicious use and safeguard academic integrity. This may also involve raising awareness of the positive and negative impacts of AI. Furthermore, given the rapid advancements in the field of artificial intelligence, it is crucial that these common policies and guidelines are kept up to date with the latest developments.

Below are some key highlights from a literature review and an analysis of GenAI chatbot mentions in the scientific literature that underpin the above recommendation:

- **Rapid growth in GenAI mentions:** Since the launch of OpenAI's ChatGPT in November 2022, the number of documents with GenAI chatbot mentions increased nearly 13-fold from November 2022 to December 2023, with OpenAI driving the uptake.
- **Concentration in specific domains:** The use of GenAI chatbots is primarily concentrated in a few research domains, with the Applied Sciences, particularly ICT, showing the highest prevalence. Health Sciences and certain subfields within the Economic & Social Sciences and Arts & Humanities also exhibit significant mentions.
- **Substantial benefits:** The use of GenAI has substantial benefits across many areas and sectors of society. In research, it notably reduces costs through more efficient literature review, data mining, data processing and analysis, and assistance in drafting manuscripts.
- **Predominant use cases:** Writing and application are the main use cases associated with GenAI chatbot mentions in scientific documents, each accounting for roughly a third of such documents. This suggests that GenAI is becoming a pervasive tool in academic writing and research applications. The distribution of documents with GenAI chatbot mentions across use cases varied by research area, document type, and document section where the mentions occurred. Notably, mentions in the acknowledgement section of documents more frequently disclose the use of GenAI in academic writing compared to other sections.
- **Strain on quality assurance and trust:** Researchers warn that increasing AI use in research and academia could strain quality assurance and raise trust issues, potentially causing societal “future shock.” While our analysis showed increased trends in the use of GenAI in research, it was not possible to ascertain the impact on actual research, methodologies, data collection, and results. The literature is also inconclusive on this aspect. Monitoring GenAI’s impact on research is increasingly important, even though it is not an explicit part of the European Union’s AI Act.
- **Ethical and risk awareness gaps:** Despite the rapid adoption of GenAI, the debate on risk awareness, ethical implications, and the impact on academic integrity is less well developed compared to discussions on the applicability and usability of these tools. For instance, the share of documents with chatbot mentions that are concerned with the ethics of GenAI usage (8%) is far below the collective shares of documents with mentions concerned with writing, application, evaluation, and development (close to 95%; note the sum across use cases

adds up to more than 100%). This highlights the need for more comprehensive discussions and policies addressing these concerns.

2. Background

Generative AI (GenAI) is defined as any artificial intelligence (AI) technology able to create new content in response to user prompts. This is achieved using various AI models (e.g., large language models, autoencoders) that learn patterns from existing data through advanced techniques like deep learning and neural networks. These models often produce outputs that are indistinguishable from human-created content.

The use of GenAI may benefit industry and society across various areas and sectors by improving efficiency and stimulating innovation. With its ability to create new content, such as text, images, music, code, and art, GenAI empowers creative professionals to explore new ideas and concepts with greater ease and speed. GenAI models can notably accelerate the prototyping process in the fashion, architecture, and product design industries by generating numerous design variations in a short amount of time. In the entertainment and gaming industry, GenAI may assist in creating realistic special effects, animations, and audio enhancements, as well as generating complex environments and enhancing character and user interactions. By automating time-consuming routine and repetitive tasks, GenAI also promises improved efficiency, productivity and cost reduction to more easily scale up such processes. By freeing up human resources and shortening time needed for processes, GenAI can lead to significant cost saving.

It is also anticipated that GenAI will be increasingly used by researchers to generate and/or analyse data, as well as to assist in drafting or reviewing manuscripts. In response to this increased usage, publishers, research institutes and governmental bodies have felt a need to introduce internal guidelines and policies to guide the responsible use of GenAI and mitigate any adverse effects.

In this policy brief, team members from Science-Metrix, NIFU and UNU-MERIT (Maastricht University) first conducted a literature review of the current debate on the use of GenAI in research, specifically on the use of chatbots (Section 2). The review is organised around the concerns raised by the use of GenAI and the debate on how to address these issues. It notably includes a review of publishers' policies and guidelines focusing on how they restrain the use of GenAI in research, as well as help ensure research integrity. Section 3 explores trends in the use of GenAI chatbots in research using documents indexed in Scopus that mentioned the use of chatbots. Section 3 also analyses the context within which these chatbots were mentioned and provides an early analysis of the imprints left by the use of GenAI chatbots in academic writing. Finally, Section 4 discusses the study findings.

3. Literature review

The idea of artificial intelligence is not recent, nor are its first implementations. Recommender systems and pattern recognition, soon followed by neural networks and deep learning models, have been in existence for more than two decades. More recently, we have witnessed an exponential growth in GenAI development with the advent of large language models (LLMs) and generative pre-trained transformers (GPTs). But, even though Stephen Hawking, Elon Musk and others published an open letter calling for more research on the societal impact of AI in 2015 (Hawking et al.), it is proving difficult to measure the impact and implications of AI use on society. This is likely due to the rapid pace of developments and the increasingly widespread

application of the technology. This statement is supported by the analysis in Section 2 which clearly points at the exponential rise in GenAI use in academic writing.

Most articles and reviews discussing the impact and implications of AI in research are concentrated in a few key areas. These areas are primarily related to health and medicine, with some coverage in business and accounting-related areas. Besides these, there is significant focus on the impact of AI on education, particularly on the implications of its use by students. At the educational level, AI poses challenges such as the potential weakening of critical thinking among students, as discussed in Section 2.3. Furthermore, the rapid advances in LLM development raise questions about the future viability of our current student evaluation system (Li et al., 2023).

Regarding the impact of GenAI, a number of socio-economic aspects are at the forefront of the societal debate. The European Parliament (2020)[1] has translated these into the following impact areas of AI on the economy and society at large:

- Impact on society: the labour market, inequality, privacy, human rights and dignity, bias, democracy
- Impact on human psychology: relationships, personhood
- Impact on the financial system: risk management, accuracy[2]
- Impact on the legal system: criminal law, tort law
- Impact on the environment and the planet: use of natural resources, pollution and waste, energy concerns, ways AI could help the planet
- Impact on trust: why trust is important, fairness, transparency, accountability, control

The European Union's AI Act (2023),[3] as part of the EU's Digital Future strategy, follows a risk-based approach to mitigate and minimise the negative impacts mentioned above. The regulatory framework of the AI Act categorises AI applications across four risk levels: minimal risk, limited risk, high risk, and unacceptable risk. The AI systems at the top risk level will be considered a 'threat to the safety, livelihoods and rights of people' (European Union, 2023) and will be banned.

Furthermore, the AI Act seeks to introduce 'transparency obligations' (European Union, 2023) for all general-purpose AI models. These obligations should allow for a better understanding and risk management of the AI models. The obligations include self-assessment, testing and model evaluations, as well as the mitigation of systemic risks, incident reporting and cybersecurity requirements.

From the six impact areas formulated by the European Parliament in 2020, the impact on research and development activities itself is missing. This lacuna, however, has been remedied by the recently published 'Living Guidelines on the Responsible Use of Generative AI in Research'[4] (European Commission, 2024) which details a number of guidelines researchers are recommended to follow, in conjunction with the European Code of Conduct for Research Integrity,[5] and the guidelines on Trustworthy AI, developed by the High-Level Expert Group on

[1]European Parliament. (2020). The ethics of artificial intelligence: Issues and initiatives (PE 634.452). EPRS | European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

[2] European Commission (2024). AI in Finance. https://finance.ec.europa.eu/news/ai-finance-2024-06-19_en

[3]https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[4]European Commission (2025). Living guidelines on the responsible use of generative AI in research https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en

[5]All European Academies (2023). European Code of Conduct for Research Integrity http://www.doi.org/10.26356/ECOC

AI.[6] These guidelines follow the four main ethical norms in research, namely reliability, honesty, respect and accountability.

The literature review, however, shows that the academic community had already adopted a certain level of self-assessment and regulation before these guidelines were established. This self-assessment and regulation relied primarily on the aforementioned principles of academic integrity and research ethics. The main points are summarised and discussed below. But first, let's explore the promises of GenAI for research.

3.1. The Benefits of GenAI for Research

Researchers and postgraduate students appreciate the capabilities of ChatGPT and similar models for their assistance in a variety of tasks. These include conducting literature reviews, enhancing the language, tone and style of research proposals or manuscripts, particularly benefiting non-native English speakers and improving equity in publishing, suggesting relevant journals for publication, brainstorming ideas, developing research designs, generating data and evidence, and promoting research through more accessible plain language summaries. The use of GenAI not only saves time but also enhances the quality of research outputs (Fecher et al. 2023), consequently boosting productivity and efficiency (Al-Zahrani, 2023; Morocco-Clarke, 2024). Access to information is crucial for researchers, and since the advent of ChatGPT, several literature-search tools (e.g., Scopus AI, Elicit, and Perplexity AI) have leveraged GenAI to help researchers responsibly process and summarise large amounts of trusted scientific knowledge at unprecedented speed. By enabling researchers to stay updated with the latest developments in their fields and integrate diverse sources of information, these tools are expected to see rapid growth in both usage and performance (Morocco-Clarke, 2024; Al-Zahrani, 2023; Dwivedi et al., 2023).

Besides researchers, scientific publishers can also benefit from the integration of GenAI technologies. GenAI can assist with the initial manuscript checking and screening process by evaluating completeness, research integrity, references, and formatting before the paper reaches editors or reviewers. This streamlines the submission process, ensuring that manuscripts meet necessary standards early on. Additionally, GenAI can recommend suitable reviewers by analysing keywords and other metadata, thereby accelerating the peer review process and enhancing the neutrality and relevance of reviewer selection. By leveraging these capabilities, GenAI helps improve efficiency, reduce administrative workload, and maintain high standards in scientific publishing.

While the benefits of GenAI in research are significant, it is essential to address challenges such as ethical considerations, data privacy, and potential biases to ensure its responsible and effective use.

3.2. The Risks and Ethics of GenAI for Research

3.2.1. Risk Awareness

Fecher *et al.* (2023) pointed out that the academic quality assurance system will likely experience increasing strain due to a rise in AI-assisted outputs in terms of writing, but an increased use of AI in research development and execution can also be expected. Erduran & Levrini (2024) echoed this sentiment and additionally pointed out that the accelerated use of AI in research might not only create trust issues within academia but also in society: Erduran &

[6] European Commission (2019). Ethics guidelines for trustworthy AI https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

Levrini cited the work of Alvin Toffler (1970), explaining that such rapid advances might prove to be a 'future shock' and provoke 'technology alienation' from the general public. Ferrara (2024) was even more specific in discussing the potential misuse of AI tools in misinformation campaigns, the generation of malicious content and the bypassing of traditional security filters with the use of AI.

Ferrara's paper (2024) is much in line with the European Parliament's (2020) concerns about the erosion of trust in government and science, as described above. The article advocates for raising awareness about 'the dual-edge nature of GenAI and LLMs' and calls for a responsible use framework that would include safeguards against malicious use, much the same as advocated for by the EU's AI Act.

3.2.2. Ethical Implications

Stahl and Eke (2024) discussed the ethical issues connected with the use of GenAI chatbots such as ChatGPT, extending the discussion to scientific writing and authorship. They proposed examining the entire AI ecosystem ─ from development to use and impact ─ and demonstrated how current ethical frameworks for emerging technologies can be holistically applied to AI. Kooli (2023) followed a similar line of thought, specifying nine main points requiring attention. Next to raising awareness and ethical implications, these include adaptation, remediation/mitigation (anti-cheating) as well as mapping future research directions including both quantitative and qualitative studies on AI impact and implications. Ghandour, Woodford and Abusaimeh (2024) applied Laudon and Laudon's (2020) five 'moral' dimensions to the use of GenAI:

- Information rights and obligations – individuals should be able to control their personal data and organisations are obligated to keep information secure and respect individual rights
- Property rights and obligations – complex challenges around IP rights, copyright, ownership and control arise when GenAI is involved in the creation of intellectual property
- Accountability and control – AI decision-making must be explainable and transparent, and there must be legal frameworks for liability and oversight
- System quality – high quality AI is essential to prevent malfunctions or 'hallucinations' that create misinformation and provide harmful advice
- Quality of life – AI must not jeopardise human values, institutions, and cultural practices, or physical health and well-being, social relationships, and autonomy

These five dimensions align relatively well with the six impact areas as defined by the European Parliament (2020).

Ghandour, Woodford and Abusaimeh (2024) also placed a disclaimer that warns of new ethical issues that might appear due to the rapid pace of technological advancement. They then, as Kooli does, call for constant 'vigilance' and for continued research to monitor the state of the art in AI development and usage, and reassess current guidelines and policies.

3.2.2.1. Impact on Academic Integrity

The issue of academic integrity seems to be closely linked with awareness of the risks of using GenAI tools in scientific research. Fecher et al. (2023) already arrived at this position by pointing out that AI generated (misleading) content raised 'concerns about scientific integrity and the nature of evidence'. Society's views of AI and the impact AI has on a socio-economic level (from misinformation to employment security) should force the scientific community to engage in a public discussion of the ethics related to these technologies and safeguard academic integrity while developing and using AI tools. As stated by Fecher, the responsibility for this aspect should rest solely with the academic community. Authors also suggested that researchers should reflect on the use of AI in the same way they report the use of other methods of data collection

and manipulation in their methodological framework. Explicit AI impact statements in publications could offer a means for this reflection (Liu et al*.,* 2022).

Though the positive benefit and assistance of GenAI tools was noted, there are also concerns about potential impacts on academic integrity. Many noted that the use of AI tools could promote plagiarism and academic dishonesty (Sullivan et al, 2023; Stokel-Walker, 2023; Jarrah et al, 2023; David, 2023; Cotton et al*.,* 2023; Eke, 2023). There were also worries regarding the authorship and ownership of content generated by AI. As pointed out by Dwivedi et al. (2023), the possibility of publishing a paper with a GenAI chatbot as a co-author raises concerns about the 'legitimacy of scholarly research' and many other problems such as privacy, misuse and lack of transparency (Dwivedi et al., 2023, p.33). Jarrah et al. (2023) stress that submitting AI-generated content without disclosure may misrepresent the author's own understanding and original input, which can be seen as deceptive. To ensure academic integrity, various studies emphasise the necessity of developing guidelines to address the use of AI tools in research, with clear regulations on authorship, data management, and transparency (Al-Zahrani, 2023; Cotton et al., 2023; Eke, 2023). See further details in Section 2.6.

3.2.2.2. Publisher Policies on the Use of GenAI in Research

As already mentioned in Section 2.1, GenAI can be applied to research itself as well as assisting in generating the outputs of this research. For the latter, both academic organisations and publishers are establishing policies for responsible use of AI in academic writing. The majority of these policies are not aimed at discouraging GenAI use, but at disclosing and delimiting the use of AI for scholarly writing. Disclosure policy should be straightforward and in line with the earlier mentioned policies on research integrity, where researchers and authors are encouraged to disclose the use of AI in their research, and the purpose of this usage in the acknowledgement, or a disclaimer, added to the paper to improve accountability, transparency and reproducibility. However, disclosure can only be effective if authors clearly understand what constitutes acceptable use of GenAI (Staiman, 2024a).

Almost all papers reviewed for this section emphasise the paramount importance of establishing proper, common definitions for fraudulent and legitimate use of AI, next to more standardized and actionable guidelines on the disclosure and limitations of AI usage (Grimaldi & Ehrler, 2023; Lin, 2024, etc.). Specifically, the reviews by Perkins and Roe (2023), Ganjavi, Eppler, Pekcan, et al. (2024) and by Lin (2024) are of interest here as they compare a large number of publisher policies on the use of GenAI. They find that the most pressing areas of improvement can be condensed into simple 'Who, What, Where' questions:

- Can AI be seen as an author? A common definition of authorship is called for; clarifying authorship credits and copyright, and who bears responsibility for quality control, proper referencing, and plagiarism (COPE, 2023, Perkins & Roe, 2024). The Committee on Publication Ethics states that 'AI tools cannot meet the requirements for authorship as they cannot take responsibility for the submitted work. As non-legal entities, they cannot assert the presence or absence of conflicts of interest nor manage copyright and license agreements.'[7]
- Which AI was used, and what was the AI's contribution to the content creation? The main issue here is the lack of standardized guidelines on 'what to report' under this heading (Lin, 2024). Alternatively, publishers could develop a list of GenAI technologies for which usage is approved, thereby removing 'the need for declaration' (Staiman, 2024a). Such a list could be modelled after the four levels of risk outlined in the earlier discussed EU AI Act (Staiman, 2024b).

[7]COPE, Authorship and AI tools, https://publicationethics.org/cope-position-statements/ai-author

- Where to report the disclosure statement? Clearer, or even standardized, guidelines could specify that the disclosure statement should be included in cover letters, methods sections, abstracts, acknowledgments, disclaimers or specific statements at the end of the article.

The main reason for the issues with these otherwise straightforward questions is the lack of detailed instructions and ambiguity in author and reviewer guidelines and publisher policies (Ganjavi, Eppler, Pekcan, et al. (2024). Additionally, there are diverging views on the scope of AI usage allowed across publishers, ranging from disallowing any AI usage to providing no guidance at all (Perkins & Roe, 2024). Two issues commonly overlooked or missing in existing guidelines and policies relate to the level of AI usage (e.g. grammar checking, editing, rewriting, writing) and the safeguarding of reproducibility. The latter goes beyond merely disclosing which AI was used by also including the prompts used and any manipulations of the algorithm and training data.

An indirectly related issue that simultaneously needs to be addressed is the use of GenAI by reviewers. Current publisher policies are mainly directed at authors, while reviewers might need similar guidance. Although some publishers have reviewer policies aimed at the use of GenAI, some do not. **From this literature review, it is clear that an industry-wide standard is needed to provide a common definition of GenAI, clear guidelines on the areas in which it can be responsibly used, and a standardised set of actionable instructions addressing the 'Who, What and Where' questions, along with the level of GenAI usage and the safeguarding of reproducibility**.

3.3. Cross-impact on Education

AI impact on research cannot be decoupled from the other main purpose of academia which is teaching. There are possibly two conflicting forces at work.

On the one hand, GenAI models such as ChatGPT are valuable tools to enhance teaching and learning by developing educational materials, providing instant feedback to students, and supporting personalised learning experiences (First, 2023). Kovanovic (2022) suggested that, similarly to other technologies in history, AI tools will be incorporated into the education system. It was also acknowledged by scholars that proper integration of AI could enhance learning and assist students in developing certain skills (Sullivan et al., 2023; Dwivedi et al., 2023; First, 2023; García-Peñalvo, 2023; Gustilo et al., 2024). On the other hand, AI tools can also have a negative impact on education and learning. Students may become over-dependent on AI tools, which could lead to a decline in critical thinking and a range of other skills (Gustilo et al., 2024; García-Peñalvo, 2023).

To guide proper use of AI in science, policymakers and educators need to establish clear guidelines and regulations to prevent the above-mentioned problems, maintain academic integrity and foster critical thinking (Sullivan et al., 2023; Schäfer, 2023; García-Peñalvo, 2023). There is also a need for teachers to adjust their teaching methods to incorporate AI effectively to prevent misuse of AI by students, and to guarantee fairness of assessments (Ifenthaler et al., 2024; Sullivan et al., 2023; Jarrah et al., 2023).

It is important to acknowledge that, apart from the research, education, and publishing concerns discussed above, there is a myriad of AI-related societal, financial, legal, and business and economic issues to address. They are not discussed in this brief but will undoubtedly be part of the ongoing debate on the impact and implications of GenAI use.

4. Data Analysis

To explore recent trends in the use of GenAI chatbots in research, documents referring to GenAI chatbots were retrieved from Scopus using regex queries.[8] Documents were then classified by research area, document type, and document section in which the mentions appeared. Additionally, the context of the mentions was analysed to extract information on the main use cases of GenAI chatbots in research overall, as well as by research area, document type and document section. The following subsections present trends in the use of GenAI chatbots in research (Section 3.1), the context within which these chatbots were mentioned (Section 3.2), and an early analysis of the imprints left by the use of GenAI chatbots in academic writing (Section 3.3).

4.1. Trends in GenAI usage in Research

Between November 2022 and March 2024, nearly 24,000 documents referenced one or more GenAI chatbots. Unless stated otherwise, all Scopus-indexed documents mentioning GenAI chatbots were included in the analysis for this section. It is possible for documents to mention multiple chatbots, multiple times, and across different sections, but each document was counted only once.

Figure 1 presents trends in the evolution of GenAI chatbot mentions from November 2022 (the launch of OpenAI ChatGPT) to December 2023.

From November 2022 (the launch of OpenAI ChatGPT) to December 2023, GenAI chatbot mentions grew rapidly in the research literature. The number of unique documents mentioning GenAI Chatbots increased from 144 (288 correcting for the likely underestimation) in November 2022 to 1,848 (3,696 after correction) in December 2023 (Figure 1, top panel). OpenAI dominated the signal, as demonstrated by shares of documents with OpenAI mentions always higher than 50%. Moreover, this share increased and capped between September and December 2023, when close to 85% of documents referred to a model in the ChatGPT family. HuggingFace models were very prominent in late 2022 and early 2023, accounting for close to half of all mentions, but their share declined near 10% in December 2023, corresponding to a roughly equivalent increase for OpenAI. Models from Meta, Microsoft, Google and others[9] remained low. It is nevertheless worth noting the slight increase for Google, which is now on par with Microsoft, both of them following closely behind Meta.

[8] Regex queries are sequences of characters used to search for specific patterns within content. In this project, the character sequences represented the names of chatbots. The queries were designed to be non-restrictive, encompassing all name variants in referring to chatbots.

[9] Anthropic, Tsinghua, Nomic AI, MosaicML, Stability AI, LMSYS, Cohere, OpenChat, Databricks, Mistral, NousResearch, DeepSeek AI, Stanford, TII.

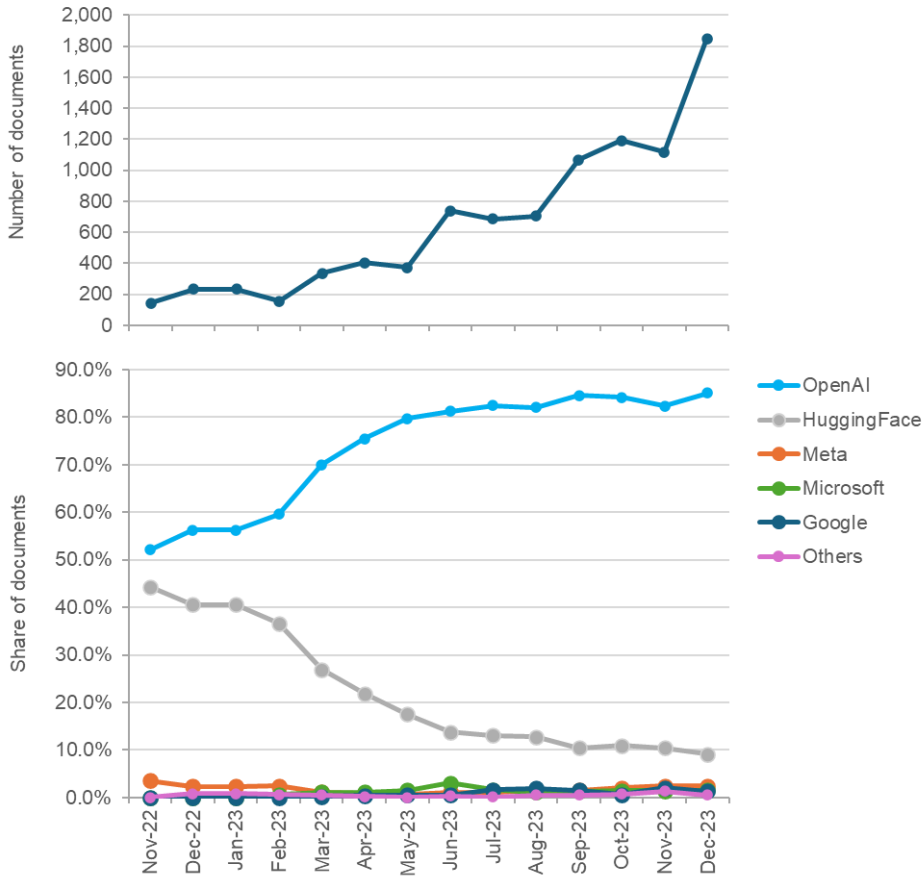


Figure 1. Number of documents referencing at least one pre-identified GenAI chatbot (upper panel) and distribution by developer (lower panel), Nov 2022–Dec 2023.

Note: Documents can mention more than one chatbot, and each chatbot can be mentioned in more than one document section. The top panel presents counts of unique documents, which are known to be underestimated by half (see next paragraph). Shares across developers in the bottom panel were computed to sum up to 100%. Documents published in 2024 were excluded due to incomplete data at the time of analysis. Additionally, close to 50% of documents lacked information on publication month, defaulting to January and causing a sharp spike in the trends for January. To address this, January 2023 data were imputed using information from the previous month (December 2022). Assuming documents with an unknown publication month are randomly distributed, the raw counts of documents in the top panel of Figure 1 should therefore be twice as large as they are. To avoid this measurement bias, the lower panel of Figure 1 presents trends in the share of documents mentioning at least one GenAI chatbot by developer, out of the sum of counts across all developers. This ensures that the shares across developers sum to 100%.

4.2. Context of GenAI Chatbot Mentions

Understanding the context within which GenAI chatbots have been mentioned in the scientific literature can offer valuable insights into their main application areas and use cases in research. To support such an understanding, documents with GenAI chatbot mentions were classified by research area (using Science-Metrix classification of science), document type (e.g., articles, conference papers, reviews, editorials, letters and notes) and document section (i.e., references, title abstracts and keywords, main publication body, and acknowledgements) in which the mentions appeared. Additionally, the context of the mentions was analysed to extract information on the main use cases of GenAI chatbots in research overall, as well as by research area, document type and document section.

This was achieved by screening and coding extracts of the GenAI chatbot mentions from more than 5,000 documents containing such mentions in their meta-matter (title, abstract, and author keywords), main document body, and/or acknowledgement section. Note that although mentions from the reference section of documents were the most common, they generally did not provide useful information to classify mentions by use case. Most mentions in the reference section of a document only provided URL links to a specific chatbot. The coding of the publication extracts by the following use cases was achieved by engineering a prompt using a Llama 3 model from Meta with 70 billion parameters:

- **Application**: Using LLM GenAI chatbots models such as GPT-2, GPT-3 for classification, translation, generating word embeddings, text mining and capturing semantic and syntactic information in local sequences of consecutive words.
- **Development**: Creating, designing, building, or improving significantly the architecture or functionality of a GenAI model to perform new tasks or achieve specific goals, excluding understanding or learning about how to develop new models.
- **Ethics**: Possible ethical issues and risks of using GenAI models.
- **Evaluation**: Assessing the performance, accuracy, or limitations of a GenAI model, including comparing its results to benchmarks, humans, or other models, or analysing its errors, biases, or robustness.
- **Writing**: Using GenAI models to assist users in drafting texts. Drafting text includes coding.
- **Other**: None of the above.

A sample of those document was also coded manually to validate the model's results. According to this manual validation, the model correctly classified 89% of cases, with 8% of cases being ambiguous, and only 4% being misclassified.

4.2.1. Documents with GenAI chatbot mentions by research area and their use cases

Documents with GenAI chatbot mentions were not evenly distributed across research domains (Figure 2, top left). Documents classified in the domain of Applied Sciences alone constituted more than half of all documents with mentions (54.4%)—a domain prone to further develop GenAI chatbot technologies relative to other domains. Besides the Applied Sciences, an appreciable share of mentions occurred in the Health Sciences (21.2%) and in Economic & Social Sciences (14.4%) (Figure 2), two areas of high relevance for the application of GenAI chatbots in research.[10] Across all domains, the most represented field was by far Information & Communication Technologies (ICT, 44.1%) from the Applied Sciences domain (Figure 2, top right). This is not surprising, as ICT is the core area from which new GenAI developments, including new applications, would naturally be expected to originate. This is further confirmed by its relative ratio of occurrence, showing that chatbot mentions in ICT occur nearly five times as frequently as would be expected if they were randomly distributed across research areas. Within the Applied Sciences, Engineering and Enabling & Strategic Technologies also featured among the seven most represented fields based on their share of documents with chatbot mentions (4.6% and 4.2%, respectively, Figure 2, top right).

To identify smaller research areas with high interest in chatbots, it is helpful to examine the occurrence of documents with chatbot mentions by field or subfield relative to all Scopus

[10]SAS, Generative AI: What it is and why it matters.
https://www.sas.com/en_ca/insights/analytics/generative-ai.html

documents in the corresponding field or subfield. A relative ratio of occurrence greater than 1 indicates a higher level of interest than expected under the assumption of a random distribution of chatbot mentions across research areas. For example, it was shown that the fields of Engineering and Enabling & Strategic Technologies were among the most represented fields based on the share of documents with mention. However, accounting for their overall size in Scopus, their relative ratio of occurrence was less than expected (0.5 and 0.4, respectively). Among the remaining most represented fields, the Social Sciences (2.2) and Economics & Business (1.2) fields, both from the Economics and Social Sciences domain, as well as Psychology & Cognitive Sciences (1.5) from the Health Sciences domain, stood out with relative ratio of occurrence above expectation.

At the subfield level, 14 categories had twice as many documents with chatbot mentions as expected (i.e., with relative ratios of occurrence ≥ 2.0, Figure 2, bottom panel), highlighting areas from four of the five domains: Applied Sciences (AS) with 5 categories, Economic & Social Sciences (ESS) with 4, Health Sciences (HS) with 3, and Arts & Humanities (AH) with 2. Only, the Natural Sciences were not represented in this group. Information & Library Sciences (ESS) led by a significant margin, with a relative ratio of occurrence 9 times as large as expected. It was followed by Medical Informatics (AS), Human Factors (HS), Artificial Intelligence & Image Processing (AS), Information Systems (AS), Software Engineering (AS), and Applied Ethics (AH), all of which had relative ratios of occurrence higher than 5.

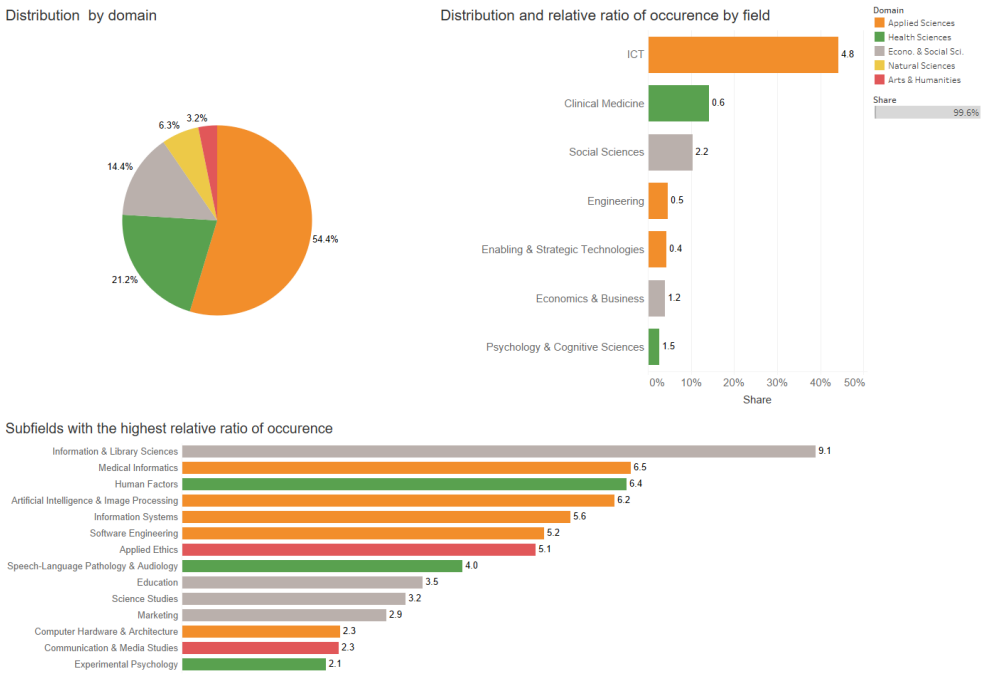


Figure 2. Distribution of documents with GenAI chatbot mentions by domain (top left) and field (top right), as well as relative distribution of such documents by subfield (bottom panel).
Note: Only fields with the highest document share are presented. Science-Metrix classification of papers by domain, field and subfield is mutually exclusive such that the sum across categories adds up to 100% (note the unclassified are not presented). The numbers next to the bars at field level correspond to the relative ratio of occurrence.

Within the Applied Sciences domain, 14.1% of documents with mentions were related to developing GenAI technologies (Figure 3). Other important use cases in the Applied Sciences

included application, writing, and evaluation (36.9%, 24.3% and 21.4%, respectively), which were also common in other domains. In fact, these latter use cases were the most common across all domains. Approximately one-third of the documents with GenAI chatbot mentions related to drafting assistance (33.7%), while another third was related to application (32.3%). About half of the remaining third referred to the evaluation of chatbot models (18.7%). As detailed below, there are nevertheless further distinctive patterns in the distribution of use cases across domains.

However, while application (36.1%) and writing (23%) are the most prevalent use cases in the Economic & Social Sciences, in similar proportions as in the Applied Sciences, the opposite was observed for the Health Sciences, with writing (47.4%) being significantly more common than application (26.2%).Perhaps the most distinctive, yet not surprising, feature of the Economic & Social Sciences is the strong relative emphasis placed on assessing the ethical issues and risks associated with the use of GenAI chatbots (18.8% versus 8% across all domains) (Figure 3). A similar, and even greater, emphasis on ethics (28.1%) was only observed in the Arts & Humanities, the domain with the smallest share of documents with chatbot mentions (3.4%). Application use cases were as frequent as ethics in this domain and were followed closely by writing (23.7%). The Natural Sciences domain was the only one, in addition to the Health Sciences, in which writing was the most common use case by a wide margin (58.9%). The application use case followed at 22.6% (Figure 3).

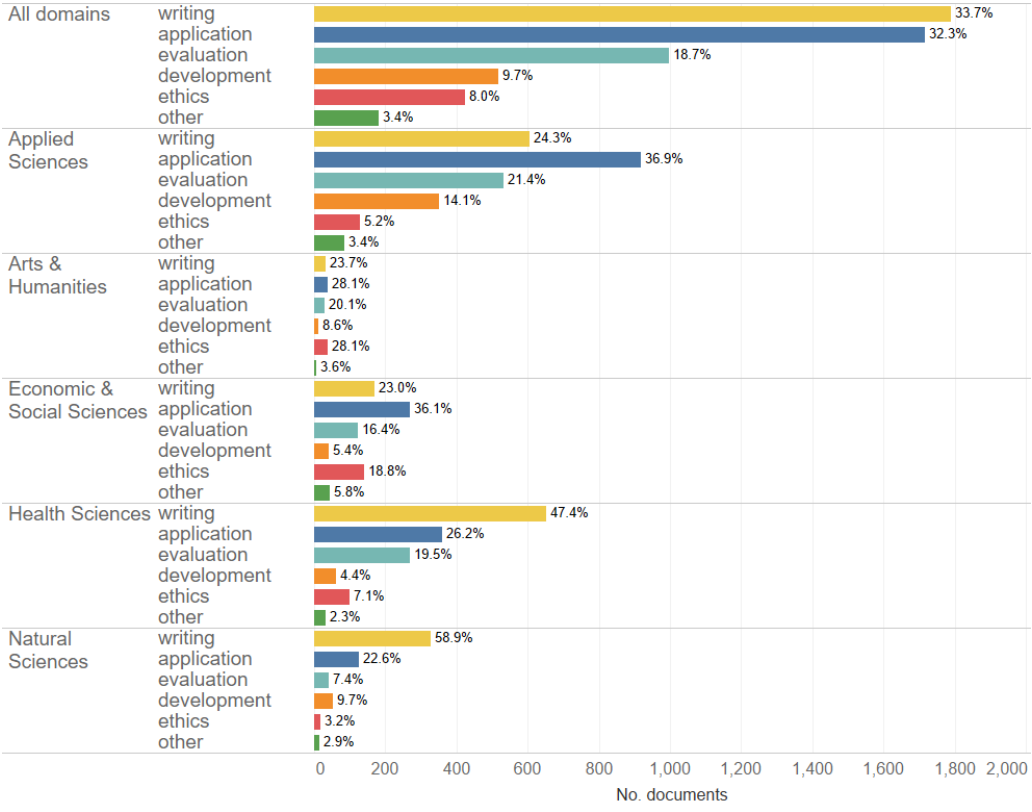
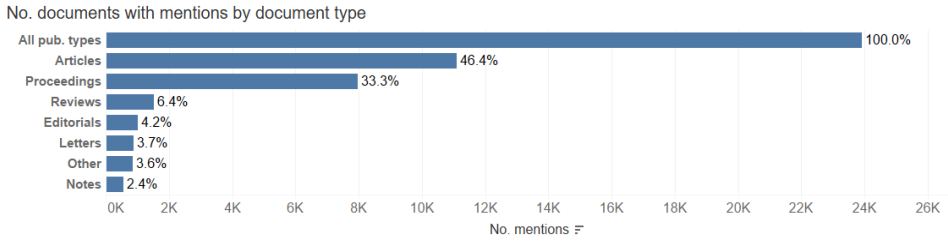


Figure 3. Context (use case) within which GenAI chatbots were mentioned by research domain
Note: The share of each use case within a given domain is provided on the right of the corresponding bar. Documents can be counted more than once if they have more than one mention and if these mentions were coded differently such that the sum across use cases can add up to more than 100%.

4.2.2. Documents with GenAI chatbot mentions by document type and section, as well as their use cases

Figure 4 shows the prevalence of GenAI chatbot mentions by document type and section. Close to 80% of documents with mentions were found in peer-reviewed documents disclosing original research, such as journal articles (46.4%) and conference papers (33.3%) (Figure 4, upper panel). While journal articles are the most represented among documents with chatbot mentions, this is primarily due to this document type being the most frequent in Scopus. When computing a relative ratio of occurrence, as was done for fields and subfields in the previous section, the prevalence of documents with chatbot mentions is below expectation for journal articles (data not shown). Conversely, it is above expectation for conference papers, which is not surprising given that this document type is most frequently used as a dissemination medium for research findings in the Applied Sciences, such as ICT and Engineering—two fields where chatbot mentions are most common— and because conference papers are often used for recent work not yet submitted to, or under review at, journals.

While mentions in Editorials amounted to less than 5% of documents (Figure 4, upper panel), their relative ratio of occurrence was also higher than expected. This highlights the interest raised by GenAI chatbots in the scientific community, particularly in discussions about their capabilities, limitations, and potential impacts in research and beyond. Letters also had a relative ratio of occurrence above expectations despite their small share (3.7%, Figure 4, upper panel). For an emerging topic with rapid development and uptake, as is observed for GenAI chatbots (Figure 1), Letters—with their concise yet peer-reviewed content—may offer a privileged medium for the timely communication of significant findings that may be of immediate interest to the scientific community.



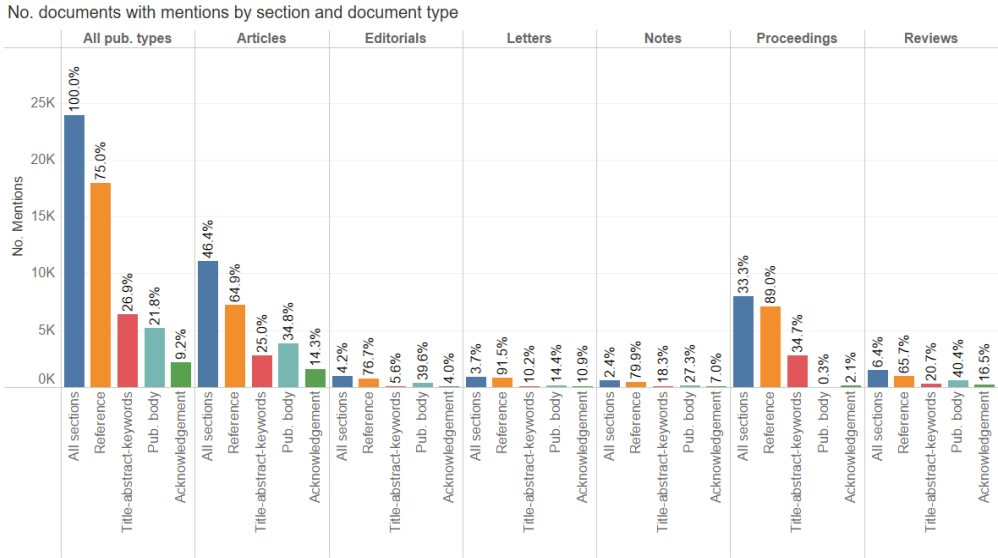


Figure 4. Distribution of documents with mentions of GenAI chatbots by document type and section.
Note: Documents can mention more than one chatbot, and each can be mentioned in more than one document section. The document type 'other' is not presented in the bottom panel.

Across all document types, the reference section had the highest occurrence of GenAI chatbot mentions (Figure 4, lower panel). Out of 24,000 documents with GenAI chatbot mentions, 75% included a mention in the reference section. Mentions in the meta-matter of documents (title, abstract, and author keywords) were found in 6,441 documents (27% of documents), while mentions in the body of the documents amounted to 5,222 documents (22% of all documents). Less than 10% of mentions were found in the acknowledgment section (2,203 documents).

When examined by document type, the proportion of mentions found within each document section varied slightly. Articles and reviews had a similar distribution of mentions across sections, with the document body (35% and 40%, respectively) and the meta-matter (25% and 21%) being most represented after the references. Interestingly, conference papers had the smallest share of mentions in the document body, at just 0.3%, and the highest share of mentions in the meta-matter (35%).

Looking at the context of chatbot mentions, articles, letters and reviews had a similar distribution across use cases, with more than 40% of the documents mentioning the use of chatbots for drafting assistance (40.1%, 56%, and 44.6%, respectively) (Figure 5). Application and evaluation were the second and third most common use cases in these document types. Editorials, like Notes, had a strong focus on ethics (35.1%). Conference papers focused more on application and evaluation (41.1% and 30.8%, respectively), with very few authors using chatbots for writing (7.9%). Relative to other document types, conference papers have a strong relative focus on development, similar to the Applied Sciences. Again, this is not surprising since conference papers are a preferred dissemination medium in that domain.

As seen in Figure 5, conference papers have, among the main document types, the highest share of mentions related to applying or developing GenAI chatbots. ICT, a field known to have a high share of conference papers among its peer-reviewed literature, is also the field with the highest share of documents mentioning GenAI chatbots (Figure 2, top right). This could be attributable to research in this field that concerns the application or development of GenAI chatbots frequently incorporating technical terminology (e.g., large language models, neural networks, generative AI) throughout its main body, while reserving the introduction of specific GenAI chatbots for the meta-matter, to convey a concise overview..

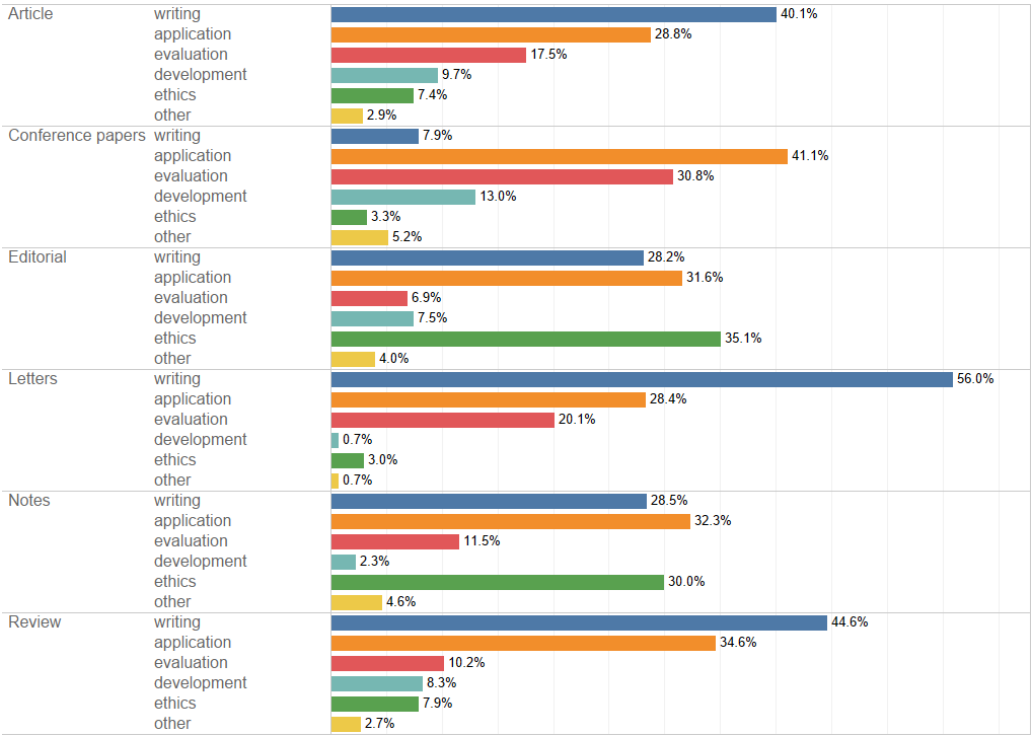


Figure 5. Context (use case) within which GenAI chatbots were mentioned by document type where those mentions were found

Note: The share of each use case within a given document type is provided on the right of the corresponding bar.

Documents can be counted more than once if they have more than one mention and if these mentions were coded differently such that the sum across use cases can add up to more than 100%.

Finally, comparing the context of mentions across document sections, a few notable variations stood out (

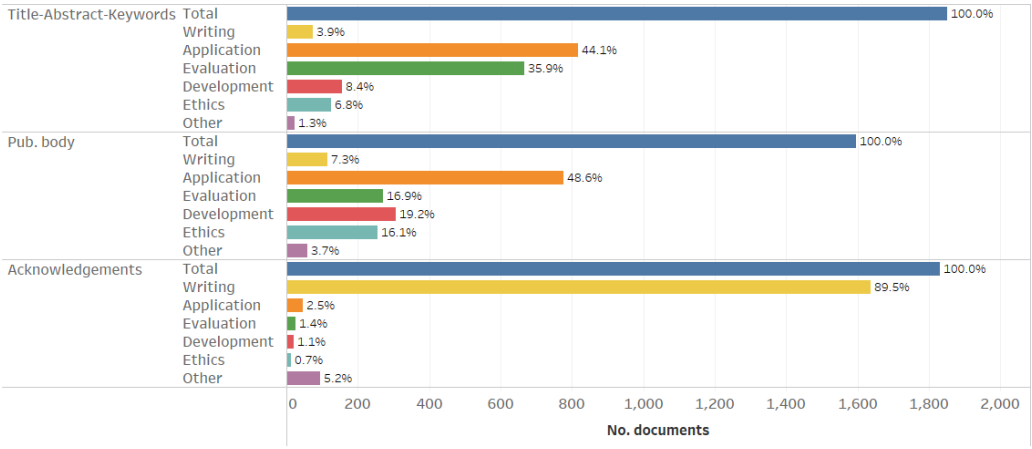


Figure 6). The title-abstract-keyword and document body sections primarily focused on applications (44.1% and 48.6%, respectively). In the title-abstract-keyword sections, the context also often revolved around evaluating the performance of existing chatbots (35.9%). In the document body, apart from applications, the context varied almost evenly between the development of new models, ethics, and evaluation (19.2%, 16.1%, and 16.9%, respectively).

Mentions in the acknowledgment section were predominantly related to acknowledging the use of GenAI for the drafting process.

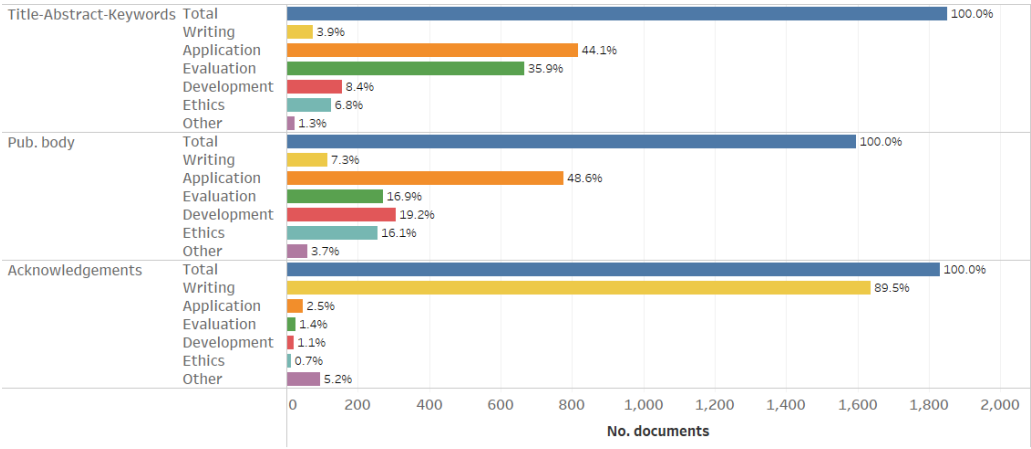


Figure 6. Context (use case) within which GenAI chatbots were mentioned by document section where those mentions were found

Note: The share of each use case within a given section is provided on the right of the corresponding bar. Documents can be counted more than once if they have more than one mention and if these mentions were coded differently such that the sum across use cases can add up to more than 100%.

4.3. Traces of GenAI Assistance in Scientific Publishing

It has been observed that certain words are favoured by large language models to a higher degree than their natural rate of use. An increased use of AI-preferred words in manuscripts could suggest the utilisation of AI chatbots for drafting manuscripts, although it is not a direct proof of usage (Gao et al., 2023).

Figure 7 shows the occurrence of selected words over three distinct 12-month periods (before, at the beginning of the use of chatbots, and one year later), with AI-preferred words shown on the left and control words on the right. AI-preferred words were selected from two lists published online.[11] Control words were chosen based on their similarity in usage to AI-preferred words. GenAI chatbots were not operational during the initial period. This analysis encompassed all papers indexed in Scopus during all three periods. An increase in the use of the words in the AI-preferred group was observed throughout the entire analysis period, with a particularly high rise noted in the third period for certain words such as 'additionally', 'explore', and 'delve'. Conversely, the control words did not exhibit an increase in mentions; rather, they displayed a relatively stable usage on average.

[11] AI-preferred words were selected from lists of words identified by two different teams (https://medium.com/@cohan_wilde/top-10-chatgpts-favorite-words-and-how-to-avoid-them-3b8b33f4a885 & https://aiphrasefinder.com/common-ai-words/)

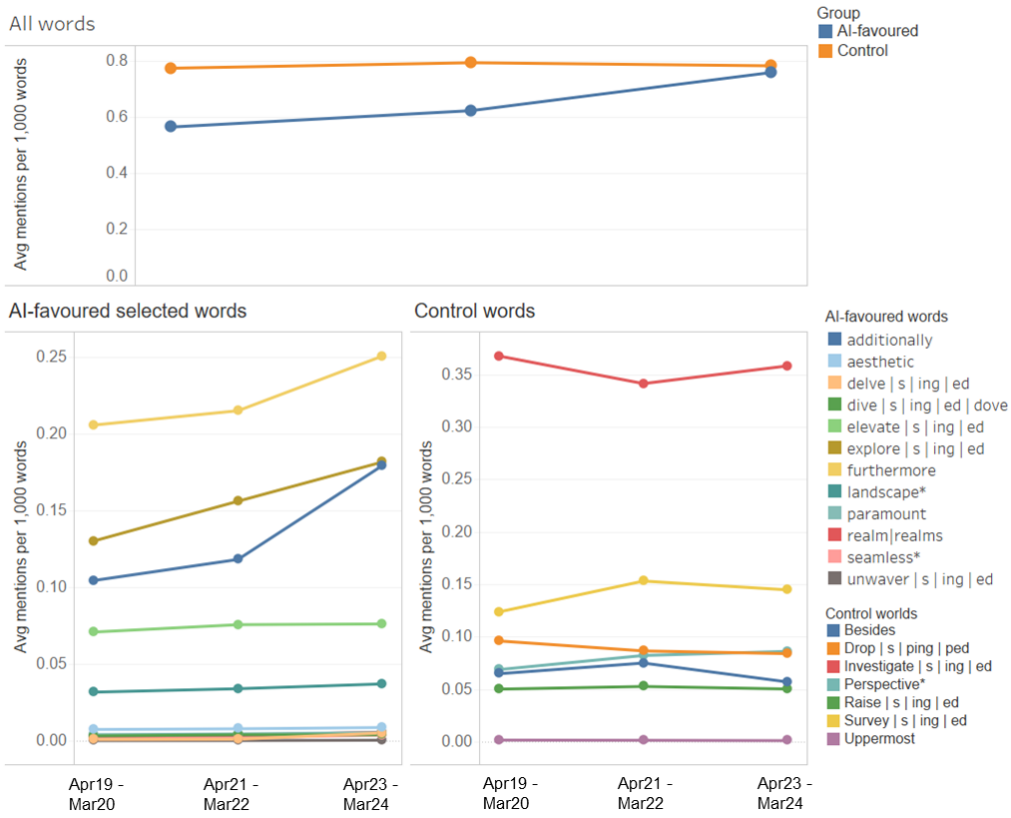


Figure 7. Mentions per 1,000 words of selected AI-favoured (left) and control (right) words over time.
Note: Words marked by an asterisk could include any additional words with the same root.

To assess the extent to which an increase in the occurrence of AI-preferred words could imply the use of GenAI chatbots, the average occurrence of AI-favoured words per thousand words was compared between documents with and without GenAI chatbot mentions for the period from April 2023 to March 2024 (Figure 8). The occurrence of AI-preferred words was higher when there was a chatbot mention in the document (in blue; 1.1 occurrence per 1,000 words versus 0.76 for the baseline, in grey). The occurrence was even greater when the mentions were found in the acknowledgement section (average of 1.25). As discussed earlier, most of the mentions found in the acknowledgement section consisted of declarations of the use of AI in the production of the manuscript.[12] Interestingly, when chatbot mentions were found in other parts of the publications (average of 1.0), the occurrence was still much higher than when there was no mention of chatbots in the document (average of 0.76).

[12] While only the average occurrence for all selected AI-favoured words was presented, the results by word were similar.

Figure 8. Occurrence of AI-favoured words in publications with and without mentions of GenAI chatbots, Apr 2022–Mar 2024
Note: * The analysis excludes documents with mentions found only in the reference section.

Further investigation is needed to determine whether the increased use of GenAI to draft manuscripts will challenge currently accepted ethical and reporting standards. Meanwhile, academic organisations and publishers have initiated responses to the rising use of GenAI to preserve the integrity of academic research. As introduced in section **Error! Reference source not found.**, such responses include the introduction of ethical policies and guidelines to establish a framework for the responsible use of GenAI in scientific publishing. While some publishers strictly forbid the use of GenAI, others allow authors to use the technology for defined purposes, provided it is properly disclosed. For example, Elsevier's policies for authors, editors and reviewers include the following:

- Authors may use GenAI to improve readability and language, but they must apply human oversight and control, disclose use of GenAI, and not list GenAI as a co-author.
- Editors and reviewers cannot upload manuscripts and peer-review reports into an AI tool, and GenAI cannot be used in the review, evaluation or decision-making process.
- GenAI cannot be used to create or alter images in manuscripts except when the use is part of the research design or method.
- The use of GenAI is prohibited for the production of artwork.

Some publishers also mention the possibility of applying forensic tools to flag possible instances of non-compliant use of GenAI technologies by content creators, such as language patterns like those detailed above or image irregularities.

5. Discussion

The primary goal of this policy brief was to provide a comprehensive overview of the current landscape and implications of Generative AI (GenAI) usage in research. By conducting a literature review and analysing trends in the use of GenAI chatbots, we aimed to synthesize the concerns and debates surrounding the responsible use of these technologies and provide early signals of their potential impacts in research. Specifically, we sought to understand how GenAI is being utilized in research, the policies and guidelines implemented by publishers to ensure research integrity, and the early impacts of GenAI on academic writing. Through this analysis, we aimed to offer insights for stakeholders, including researchers, publishers, and policymakers, to navigate the rapidly evolving landscape of GenAI in research.

Since the launch of OpenAI's ChatGPT in November 2022, the landscape of generative AI has evolved rapidly, with numerous new models and updates released at an unprecedented pace. For instance, in just the first few months of 2023 alone, six major advancements were made, including the introduction of GPT-4 and various specialized models from companies like Google, Microsoft, and Meta, reflecting contributions from multiple distinct developers across the industry (PYMNTS, 2024). Results from the analysis of the scientific literature also show staggering growth in the mention of GenAI chatbots in research manuscripts. From November 2022 to December 2023, the number of documents with GenAI chatbot mentions increased

nearly 13-fold, an exponential growth mostly driven by easily accessible models, particularly OpenAI's ChatGPT.

The literature review, as well as the analysis of GenAI chatbot mentions in scientific documents, showed that the focus of interest in GenAI in the academic world is so far concentrated in a few domains, fields and subfields. In general, the Applied Sciences domain stands out the most, which is not surprising as the development of GenAI has its roots in the field of ICT. The whole domain of Health Sciences sees many mentions of GenAI, especially if one considers the subfields of the Applied Sciences that are medicine-adjacent, such as Medical Informatics. Within the Health Sciences, some of the subfields standing out most also have a strong connection with developments taking place in ICT, such as the subfield of Human Factors. Apart from these, subfields from the domain of the Economic & Social Sciences, as well as of the Arts & Humanities, do stand out. In the former domain, the subfields of Information and Library Sciences and Science Studies, where this report would fall, are concerned with the use of GenAI. Besides them, GenAI chatbots frequently appear in Marketing and Education research, the latter subfield being concerned with teaching, one of the two main purposes of academia besides research.

The literature review also revealed the unease of educators with the use of AI in education, which can be explained by the 'future shock' effect that the rapid rise in the use of GenAI solutions, as offered for instance by OpenAI, has had on academia and society as a whole. Although the misgivings that educators seem to have with regards to students using AI assistance are understandable as they could lead to a decline in critical thinking and a range of other skills (Gustilo et al., 2024; García-Peñalvo, 2023), scholars themselves seem to be equally willing to use GenAI assistance in their writing. Across all domains, **our results show that writing is the main use case associated with GenAI chatbot mentions in scientific documents, along with application, each accounting for roughly a third of all instances**. We also found an increase in the frequency of 'AI-favoured' words in scientific documents with chatbot mentions versus those without; an increase not observed for control words. The conclusion here is that AI assistance in research and academic writing is becoming increasingly pervasive, corroborating results by Deike (2024). Additionally, our results on the pattern of 'AI-favoured' word usage as a function of the document section in which chatbot mentions appeared suggest that not all assistance in academic writing is transparently disclosed through declarations on the use of GenAI for such purposes.

In the Arts & Humanities domain, the Ethics subfield has five times as many documents mentioning GenAI chatbots as expected if chatbot mentions were randomly distributed across research areas. However, the share of documents with such mentions concerned with the ethics of GenAI usage stands at 8% across all domain. This is well below the collective shares of documents with mentions concerned with writing, application, evaluation, and development (close to 95%).[13] Given the indications offered by our data analysis, the debate on GenAI usage in research has been cautious, to say the least. Our literature review shows that while the applicability and usability of GenAI tools in areas such as the Health Sciences are researched and discussed at length, the debate on risk awareness, ethical implications and the impact on academic integrity is less well developed (Al-Zahrani, 2023; Eke, 2023; Kooli, 2023; Ferrara, 2024; Ghandour, Woodford and Abusaimeh, 2024). Even though policies addressing these concerns are being developed at the higher levels of governance, their practical implementation does not seem to have trickled down to the academic workplace yet. More work is needed in this area.

As previously noted, the development within GenAI technology, and the field of artificial intelligence as a whole, is extremely rapid. Already, proposals for fully automated full-stack GenAI implementations of the whole scientific process are surfacing (Lu et al. 2024). Although

[13] Recall that the sum of shares across use cases adds up to more than 100% as documents can relate to more than one use case.

some key stakeholders have already established guidelines to regulate research practices as relates to the use of GenAI, notably in academic writing, it remains of utmost importance for the academic community and policy makers, both at the national and supranational levels, to develop and adopt common policies and guidelines regarding the integration of these technologies into scientific practice and reporting workflows to prevent malicious use and safeguard academic integrity (Ferrara, 2024). Furthermore, acknowledging the rapid pace of developments in the field of artificial intelligence, it is crucial that these common policies and guidelines are kept up to date with the latest advancements (Ghandour, Woodford and Abusaimeh, 2024).

6. Bibliography

Al-Zahrani, A. M. (2023). The impact of generative AI tools on researchers and research: Implications for academia in higher education. *Innovations in Education and Teaching International*. https://doi.org/10.1080/14703297.2023.2271445

Ciaccio, E. J. (2024). Detecting ChatGPT in published documents: Chatbot catchphrases and buzzwords. *Informatics in Medicine Unlocked*, https://doi.org/10.1016/j.imu.2024.101516

COPE (2023). *Artificial intelligence and authorship*. Retrieved 06/08/2024 from: https://publicationethics.org/news/artificial-intelligence-and-authorship

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. https://doi.org/10.1080/14703297.2023.2190148

David, E. (2023). Pausing AI might be a good idea, but companies won't do it. *Business Insider*. https://www.businessinsider.com/openai-elon-musk-pause-development-letter-never-happen-2023-4

Deike, M. (2024). A "Delve" into the evidence of AI in production of academic business literature. *Journal of Business & Finance Librarianship.* A "Delve" into the evidence of AI in production of academic business literature: Journal of Business & Finance Librarianship: Vol 0, No 0 - Get Access

Dwivedi, Y. K., Kshetri, N., Hughes, L. et al. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Eke, D. (2023). ChatGPT and the rise of GenAI: Threat to academic integrity? Journal of Responsible Technology. https://doi.org/10.1016/j.jrt.2023.100060

Erduran, S., & Levrini, O. (2024). "The impact of artificial intelligence on scientific practices: an emergent area of research for science education", *International Journal of Science Education*, https://doi.org/10.1080/09500693.2024.2306604

European Commission and the High-Level Expert Group on AI (2019) *Ethics Guidelines for Trustworthy AI*. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419

European Commission (2023). *AI Act*, Retrieved 18/06/2024 from: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

European Commission (2023) *European Code of Conduct for Research Integrity, Revised Edition*. http://www.doi.org/10.26356/ECOC

European Commission (2024) *AI in Finance*. Retrieved 17/09/2024 from: https://finance.ec.europa.eu/news/ai-finance-2024-06-19_en

European Commission (2024) Living Guidelines on the Responsible Use of Generative AI in Research. *ERA Forum Stakeholders' document*. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en

European Parliament (2020). *The ethics of artificial intelligence: Issues and initiatives* (PE 634.452), European Parliamentary Research Service. Retrieved 18/06/2024 from: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf

Fecher, B., Hebing, M., Laufer, M. et al. (2023). "Friend or foe? Exploring the implications of large language models on the science system", *AI & Society*. https://doi.org/10.1007/s00146-023-01791-1

Ferrara, E. (2024). "GenAI against humanity: nefarious applications of generative artificial intelligence and large language models", *Journal of Computational Social Science*. https://doi.org/10.1007/s42001-024-00250-1

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*. https://doi.org/10.37074/jalt.2023.6.1.22

Gao, C.A., Howard, F.M., Markov, N.S. et al. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*. https://doi.org/10.1038/s41746-023-00819-6

Ganjavi C, Eppler M B, Pekcan A, Biedermann B, Abreu A, Collins G S et al. (2024) Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. *British Medical Journal*. https://doi.org/10.1136/bmj-2023-077192

García-Peñalvo, F. J. (2023). The perception of Artificial Intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society*. https://doi.org/10.14201/eks.31279

Ghandour, A., Woodford, B., and Abusaimeh, H. (2024) "Ethical Considerations in the Use of ChatGPT: An Exploration through the Lens of Five Moral Dimensions", *IEEE Access*. https://doi.org/10.1109/ACCESS.2024.3394243

Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity*. https://doi.org/10.1007/s40979-024-00153-8

Hawking, S., Musk, E., Wozniak, S. et al. (2015) *Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter*. Retrieved 19/06/2024 from: https://futureoflife.org/open-letter/ai-open-letter/

Ifenthaler, D., Majumdar, R., Gorissen, P. et al. (2024) "Artificial Intelligence in Education: Implications for Policymakers, Researchers, and Practitioners", *Technology, Knowledge and Learning*, https://doi.org/10.1007/s10758-024-09747-0

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say? *Online Journal of Communication and Media Technologies*. https://doi.org/10.30935/ojcmt/13572

Kooli, C. (2023) "Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions", *Sustainability*, https://doi.org/10.3390/su15075614

Kovanovic, V. (2022). The dawn of AI has come, and its implications for education couldn't be more significant. *The Conversation* (Australia and NZ). https://www.proquest.com/newspapers/dawn-ai-has-come-implications-education-couldnt/docview/2754181122/

Laudon, K., & Laudon, J. (2020). *Management Information Systems: Managing the Digital Firm*, Harlow: Pearson Education

Li, Y., Sha, L., Yan, L., et al. (2023). Can large language models write reflectively, *Computers and Education: Artificial Intelligence.* https://doi.org/10.1016/j.caeai.2023.100140

Lin, Z. (2024) Towards an AI policy framework in scholarly publishing. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2023.12.002

Liu, D., Nanayakkara, P., Sakha, S., et al. (2022) "Examining Responsibility and Deliberation in AI Impact Statements and Ethics Reviews", *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, https://doi.org/10.1145/3514094.3534155

Lu, C., Lu, C., Lange, R. T., Foerster, J., Clune, J., & Ha, D. (2024). The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. *arXiv preprint arXiv:2408.06292*.

Morocco-Clarke, A., Sodangi, F. A., & Momodu, F. (2024). The implications and effects of ChatGPT on academic scholarship and authorship: a death knell for original academic publications? *Information and Communications Technology Law*. https://doi.org/10.1080/13600834.2023.2239623

Perkins, M., & Roe, J. (2024) Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis*. F1000Research*. https://doi.org/10.12688/f1000research.142411.2

PYMNTS (2024) GenAI's Pace of Development Is Shattering Moore's Law. https://www.pymnts.com/artificial-intelligence-2/2024/genais-pace-of-development-is-shattering-moores-law/

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. *Journal of Science Communication*. https://jcom.sissa.it/article/pubid/JCOM_2202_2023_Y02

Stahl, B., & Eke, D. (2024) "The ethics of ChatGPT – Exploring the ethical issues of an emerging technology", *International Journal of Information Management*, https://doi.org/10.1016/j.ijinfomgt.2023.102700

Staiman, A. (2024a). *Dark Matter: What's missing from publishers' policies on AI generative writing?* Retrieved 01/08/2024 from: https://www.digital-science.com/tldr/article/dark-matter-whats-missing-from-publishers-policies-on-ai-generative-writing/

Staiman, A (2024b). *Woefully Insufficient Publisher Policies on Author AI Use Put Research Integrity at Risk*. Retrieved 01/08/2024 from: https://scholarlykitchen.sspnet.org/2024/07/22/woefully-insufficient-publisher-policies-on-author-ai-use-put-research-integrity-at-risk/

Stokel-Walker, C. J. N. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*. https://doi.org/10.1038/d41586-023- 00107

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*. https://doi.org/10.37074/jalt.2023.6.1.17

Toffler, A. (1970). *Future Shock*. New York: Random House.

7. Annex – Study Methodology

7.1. Declaration of generative AI and AI-assisted technologies in the working process

The authors acknowledge the use of GenAI chatbots for data mining in the literature review (Section 2), for the coding the use cases in the data analysis section (Section 3), and for improving the readability and language of the text. The authors carefully reviewed and edited the proposed changes as needed and take full responsibility for the content of the publication, including the critical analysis and conclusions they developed.

7.2. Literature Review

Scopus citation index[14] was screened to identify articles discussing both the impact and the implications of GenAI for research. Papers of interest consist of the combination of measurement of AI use IN research and the effects of this use ON research. This resulted in the following Boolean query:

- (TITLE-ABS-KEY(artificial intelligence) OR TITLE-ABS-KEY("AI")) AND (TITLE-ABS-KEY(research) AND TITLE-ABS-KEY(implicat*)) AND (TITLE-ABS-KEY(impact) OR TITLE-ABS-KEY(usage) OR TITLE-ABS-KEY(monitor*)) AND PUBYEAR > 2017 AND PUBYEAR < 2025

The search was limited to the publication years from 2018 and onwards. The corpus that resulted from this search contains 2,092 papers.

GPT-4o[15] was employed to recommend the inclusion or exclusion of papers based on a set of prompts:

"Publications discussing the impact AND implications of the use and usage of artificial intelligence in research"

"Publications discussing the impact OR implications of the use and usage of artificial intelligence in research"

These prompts had to represent both instances of impacts and/or implications, as well as benchmarking it against the other HEI's main pillar, education. The education pillar was likewise concerned with the use of AI.

As an additional benchmark, research was added to, and replaced with 'education':

"Publications discussing the impact and implications of the usage of artificial intelligence in research and education"

"Publications discussing the impact and implications of the usage of artificial intelligence in education"

Table 1 shows the results of this exercise. Using the prompt, a quarter of the corpus were considered relevant with respect to the research AND prompt. The OR prompt selection obviously contained more papers but not to the extent expected. The benchmarks were

[14] See: https://www.scopus.com/

[15] See: https://openai.com/index/hello-gpt-4o/

reassuring in that only a few were selected that, according to the AI, dealt with education related issues only.

Table 1. GPT-4o results for the inclusion or exclusion in the literature review

Prompt	Exclude	Include
research_AND	1542	550
research_OR	1209	883
research_education	1636	456
education	1829	263

Manual validation of the 550 papers selected by the AI circumscribes furthermore the publication set to publications that solely deal with the impact and implications of AI use in research. The exclusion criteria was: (i) science focus; not a niche area focus but sufficiently broad to actual discuss the impact across science as a whole (generalisability). (ii) citations; as the corpus is published in a constricted time-period of just 8 years citation counts can only be used sparingly as an exclusion criteria although obviously those papers that are highly cited should be part of the final corpus. (iii) Media types; notes, letters etc. will feature further down the list for inclusion, and (iii) Publication years; in combination with citations the publication year can be a good criterion to select for inclusion. This final step resulted in a publication set of 25 papers. This is not surprising as the field is relatively new and most papers discuss AI related issues on a specific topic, or on an otherwise non-generalisable level.

7.3. Quantitative Analysis

7.3.1. Identification of publications with mentions of GenAI chatbots

The publications and associated data in the present study were extracted from the Scopus database, owned by Elsevier. Scopus includes articles, reviews and conference papers from all fields and subfields of research. Data from Scopus used for this study were accessed in March 2024. All publication types were included in the analysis, except preprints.

The selection of the GenAI chatbot models was based on a list of the most used chatbots proposed by LMSYS (https://chat.lmsys.org/?leaderboard), as of May 1, 2024). The list was expanded to include Meta IA, Google Bard, and Microsoft Copilot. Relevant papers were retrieved using a relatively permissive list of regex queries (e.g., "Claude-, "Mistral-", "GPT", "Llama") to retrieve as many papers as possible. This led to the inclusion of models not included on the list (e.g., former GPT-2, future GPT-5, other Llama models). The use of non-restrictive keywords and the fact that some bots had very common names (e.g., Claude, Gemini, Dolphin), led to the inclusion of false positives that were excluded in a second round of regex queries. Using samples of 100 papers in each of the above-mentioned sections, the precision (e.g., proportion of retrieved papers that are relevant to GenAI chatbots) was estimated at 98% (approximately 2% false positives per section).

7.3.2. Context of use of GenAI

Extracts of publications with mentions of chatbots in the publication body, title-abstract-keywords, and acknowledgements sections were screened for the context within which these mentions were made. For each of these publication section, up to 2,000 distinct papers were examined using the GenAI model Llama-3-70b, for a total of 5,179 papers. The model was trained to select the most relevant context based on the definition provided in the prompt. Precision was estimated in samples of 100 papers. All text duplicates (the same paper section extract retrieved using different keywords) coded differently were manually checked and the best context was selected (or corrected). Some paper extracts could have more than one context.

Prompt: *Should "application" be defined as using LLM genAI models like GPT-2, GPT-3 for classification, translation, generating word embeddings, text mining and capturing semantic and syntactic information in local sequences of consecutive words, "development" be defined as creating, designing, building, or improving significantly the architecture or functionality of a genAI model to perform new tasks or achieve specific goals, excluding understanding or learning about how developing new models, "evaluation" be defined as assessing the performance, accuracy, or limitations of a genAI model, including comparing its results to benchmarks, humans, or other models, or analyzing its errors, biases, or robustness, "ethics" as the possible ethical issues and risks of using genAI models, and "writing" as using genAI models to assist users drafting texts.*

Manual coding by use case was also made to ascertain that the coding produced by the model were sound. Only very minor differences were found between the manual coding and the AI coding (data not shown).

7.3.3. Occurrence of AI-favoured words in publications

Identifying typically GenAI-generated words was based on a list of words that tend to occur in AI-generated text at a much higher frequency than in other contexts (Ciaccio 2024). The list of AI-favoured words and control words is shown in Table 2. Those words were selected on a set of words found on aiphrasefinder.com and medium.com. [16,17] An effort was made to choose control words with similar meanings; however, the primary objective was to select control words that were not included in any online lists that could categorize them as AI-preferred terms.

[16]AI Phrase Finder, The 100 Most Common AI Words, https://aiphrasefinder.com/common-ai-words/

[17]Medium, Top 10 ChatGPT's Favorite Words. https://medium.com/@cohan_wilde/top-10-chatgpts-favorite-words-and-how-to-avoid-them-3b8b33f4a885

Table 2. Identification of AI-favoured words and control words

AI-favoured words	Control words
Delve | s | d | Delving	Investigate | s | d | Investigating
Dive | s | d | Diving | Dove	Drop | s | ped | ping
Elevate | s | d | Elevating	Raise | s | d | Raising
Additionally	Besides
Aesthetic	
Explore | s | ing | ed	Survey | s | ing | ed
Furthermore	
Landscape*	Perspective*
Paramount	Uppermost
Realm | s	
Seamless*	
Unwaver | s | ing | ed	

Note: Asterisk indicates that words can be expanded (e.g., seamless = seamless and seamlessly)*

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 (european-union.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.

The study examines the surge in GenAI chatbot mentions in scientific literature, showing a 13-fold increase from November 2022 to December 2023. The use of GenAI chatbots in scientific research is mainly in ICT and Applied Sciences, where AI improves research efficiency. Key applications include writing and practical implementation, demonstrating the tool's widespread use in academic writing and research. Nonetheless, the increasing use of AI in research and academia raises concerns about quality assurance and trust issues.

Research and Innovation policy

Publications Office
of the European Union