






Rise of the machines: trends and challenges of implementing AI in biomedical scientific writing

Michał Fornalik^{1*} , Magdalena Makuch² , Anna Lemanska¹, Sandra Moska¹, Monika Wiczewska¹, Iwona Anderko¹, Laura Stochaj¹, Marta Szczygiel¹, Aleksandra Zielińska^{3*} 

¹Faculty of Medicine, Poznan University of Medical Sciences, 61-701 Poznan, Poland

²Department of Immunology, Jagiellonian University, 30-387 Cracow, Poland

³Department of Biotechnology, Institute of Natural Fibers and Medicinal Plants National Research Institute, 60-630 Poznan, Poland

***Correspondence:** Michał Fornalik, Poznan University of Medical Sciences, 61-701 Poznan, Poland. michalfornalik.contact@gmail.com; Aleksandra Zielińska, Department of Biotechnology, Institute of Natural Fibers and Medicinal Plants National Research Institute, 60-630 Poznan, Poland. Aleksandra.zielinska@iwnirz.pl

Academic Editor: Zhaohui Gong, Ningbo University, China

Received: May 18, 2024 **Accepted:** July 23, 2024 **Published:** September 4, 2024

Cite this article: Fornalik M, Makuch M, Lemanska A, Moska S, Wiczewska M, Anderko I, et al. Rise of the machines: trends and challenges of implementing AI in biomedical scientific writing. *Explor Digit Health Technol.* 2024;2:235–48. <https://doi.org/10.37349/edht.2024.00024>

Abstract

Artificial intelligence (AI) technology is advancing significantly, with many applications already in medicine, healthcare, and biomedical research. Among these fields, the area that AI is remarkably reshaping is biomedical scientific writing. Thousands of AI-based tools can be applied at every step of the writing process, improving time effectiveness, and streamlining authors' workflow. Out of this variety, choosing the best software for a particular task may pose a challenge. While ChatGPT receives the necessary attention, other AI software should be addressed. In this review, we draw attention to a broad spectrum of AI tools to provide users with a perspective on which steps of their work can be improved. Several medical journals developed policies toward the usage of AI in writing. Even though they refer to the same technology, they differ, leaving a substantially gray area prone to abuse. To address this issue, we comprehensively discuss common ambiguities regarding AI in biomedical scientific writing, such as plagiarism, copyrights, and the obligation of reporting its implementation. In addition, this article aims to raise awareness about misconduct due to insufficient detection, lack of reporting, and unethical practices revolving around AI that might threaten unaware authors and medical society. We provide advice for authors who wish to implement AI in their daily work, emphasizing the need for transparency and the obligation together with the responsibility to maintain biomedical research credibility in the age of artificially enhanced science.

Keywords

Artificial intelligence, AI-based tools, biomedical scientific writing, medical journals, biomedical publishing



Introduction

Recent advancements in artificial intelligence (AI) technology have reshaped the boundaries of what was previously deemed as original scientific work. “Artificial intelligence” represents a vast field of technology aimed at comprehending and simulating human-like behavior by computer programs based on analyzing large amounts of data [1]. To observe, learn, understand, and react, humans utilize a complex network of neuronal wires operated by an apparatus of proteins and molecules. Every part of this machinery is encoded by DNA, which serves as a blueprint for all aspects of life. AI tries to replicate those actions by replacing DNA with digital computer programming language. Nomenclature regarding AI is changing rapidly; however, this article divides this technology into generative and non-generative AI. According to user instructions, the former can generate ‘new’ content like text, images, or movies. At the same time, the latter focuses on specific tasks with a much narrower field of action [2].

AI of both types is currently revolutionizing fields of science and medicine and can permanently transform how research is conducted and defined. Biomedical and clinical sciences rank as the second most popular fields of research involving AI, just behind information and computing sciences [3]. The application of AI in the medical field, as discussed across the newest scientific reports, presents a multifaceted landscape with numerous phenomena, advantages, and disadvantages [4]. AI increasingly permeates various facets of healthcare, from medical imaging through health insurance adjudication to prospective studies and randomized controlled trials [5, 6]. The growing interest and need for a dedicated platform to publish high-quality medical AI research have led to the introduction of new series in prestigious medical journals and even the launch of entirely new journals [7, 8]. NEJM AI, set to launch in 2024, acknowledges the vast potential and challenges of AI applications across all areas of medicine and care delivery, such as automating medical note dictation and synthesizing patient data [7].

As for today, AI has proven to be invaluable in various tasks related to biomedical research. Examples include experimental design [9], acting as a writing assistant [10], as well as recommending the most suitable journals for publishing articles and sharing results with the scientific community [11].

In the case of medical writing, it is proposed that physicians can leverage AI to generate medical reports, summarize research papers and clinical trial results, and write medical textbooks, guidelines, and articles [12]. Khalifa et al. [13] identified six main domains where implementing AI can significantly improve authors’ performance: idea development, content development, literature review, data management, editing and review, and communication.

There are currently thousands of AI software available on the market, many of which find applications in academia, particularly in scientific writing [14, 15]. It may pose a challenge to determine for which task a particular tool might be helpful. In this review, we draw attention to a broad spectrum of AI tools that can be implemented at each step of the writing process. We emphasize their capabilities that can help optimize and increase the efficiency of the creative workflow. Our summary will be valuable for authors intending to integrate AI tools into their daily practice and help them optimize their work. We will also clarify the most common obscurities the authors might struggle with while using AI tools and address several issues and ambiguities that may arise while implementing them into scientific work.

ChatGPT and other AI tools

ChatGPT is undeniably the most popular AI tool, with over 180 million users [15, 16], and has already received extensive coverage in numerous publications [17–21]. Nevertheless, it is important to mention its potential for development in the academic field, particularly medicine. Nowadays, AI can act as a personal, educational assistant for medical students [22], and even pass the tests that those students must take. For example, ChatGPT passed or came close to passing the United States Medical Licensing Exam (USMLE) [23] and presented a relatively high performance in various specialty exams including radiology [24], radiation oncology [25], and rheumatology [26]. However, it is essential to recognize that different platforms offer a broader range of utilities. This includes generative and non-generative AI tools that can be utilized in medical academic writing (Figure 1).

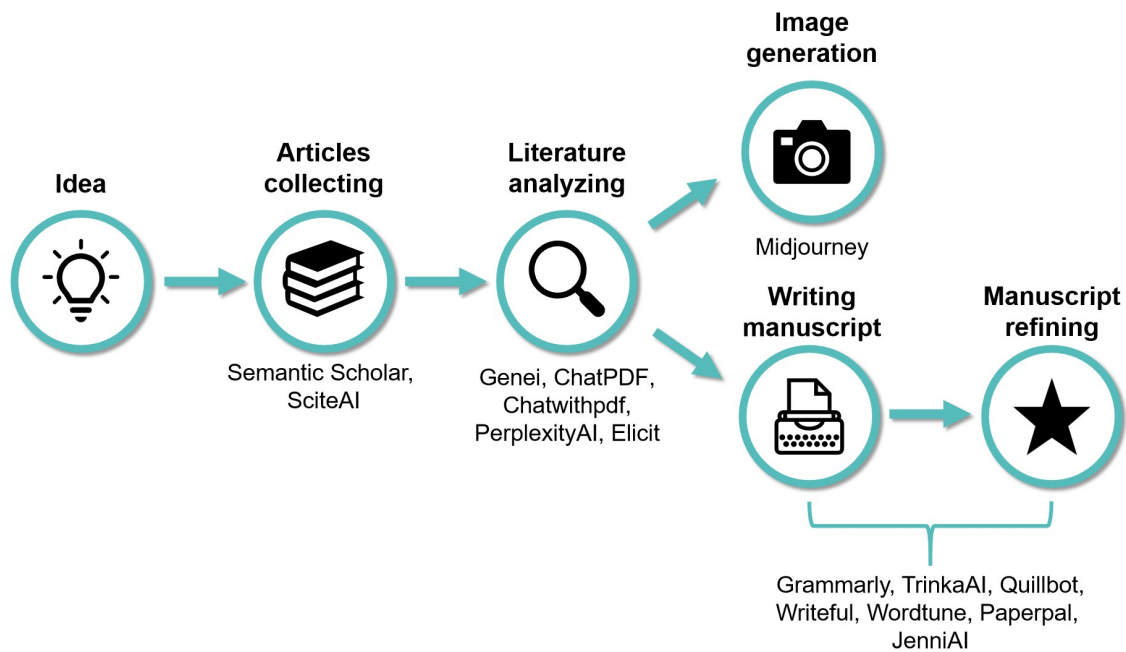


Figure 1. Visual representation of the functions offered by AI-based tools that can be applied in the writing process with examples of specific software. AI can be implemented to generate ideas, collect articles [29–31], analyze literature [30, 37, 40–43], write and refine manuscripts [30, 48–50, 53–57], and aid with image generation [82]

Several prestigious journals struggle with this rapidly developing technology [27]. The newly established editorial policies and author guidelines mainly focus on large language models (LLMs) for generative AI and AI-generated images. An analysis of publishers’ and journals’ guidelines for authors by Ganjavi et al. [27] revealed that (by the year 2023) only 24 of the top 100 most important publishers released policies toward the usage of generative AI. Most comprised statements on including generative AI as a publication co-author, almost unanimously forbidding it. 54% of them referred explicitly to ChatGPT in their policies. Numerous publishers with no or unspecific guidance threaten scientific integrity and transparency since it leaves a considerable gray area susceptible to abuse. It is crucial to establish well-refined and comprehensive policies that will include more than just LLMs as the spectrum of AI tools is becoming continually broader.

In the following sections, we describe other tools that are particularly useful in writing and may go beyond typical publishers’ or journals’ AI policies. In this task, we will follow an author’s typical organizational approach when initiating a new biomedical article. All the tools we mention in this article are summarized in Table 1.

Table 1. The list of popular AI-based tools is divided into subcategories

| Step | AI tool | Function | Reference(s) |
|----------------------|------------------|---|--------------|
| Articles collecting | Semantic Scholar | Identifying the proper papers, extracting meaning, checking highly influential citations, and recommending the latest documents that may be of interest | [29, 30] |
| | Scite.ai | Identifying the most significant papers and demonstrating the connections between them | [30, 31] |
| | PICO Portal | Identifying the most relevant papers meeting specific PICO criteria | [35] |
| Literature analyzing | Genei | Condensing research papers, extracting the essential information | [37] |
| | ChatPDF | Extracting information from PDF files by answering the posed questions (e.g., primary outcomes from the study, conclusions, limitations) | [40, 41] |
| | Chatwithpdf | | |
| | Perplexity AI | Answering the posed questions and providing resources | [30, 42] |
| | Elicit | Summarizing papers, finding themes and concepts across papers, extracting data, synthesizing the findings | [43] |
| Writing and refining | Grammarly | Context-specific suggestions to help with proper grammar, punctuation, and spelling, style improvements | [48, 49] |

Table 1. The list of popular AI-based tools is divided into subcategories (*continued*)

| Step | AI tool | Function | Reference(s) |
|--------------|--------------|--|--------------|
| a manuscript | Trinka AI | Checking the word choice, usage, style, and grammar | [30, 48] |
| | QuillBot | Paraphrasing, grammar check, ensuring the appropriate vocabulary, tone, and style, text summarizing | [30, 50] |
| | Writefull | Text revision, language and tone check, paraphrasing, title and abstract generator | [30, 53] |
| | Wordtune | Paraphrasing, eliminating grammar, and spelling errors | [54, 55] |
| | Paperpal | Ensuring adherence to stringent quality standards (specialized in academic papers) | [30, 56] |
| Peer review | Jenni AI | Autocomplete in-text citations, comprehensive prompts generation, paraphrasing, paper refining, plagiarism check | [30] |
| | Penelope.ai | Paper evaluation, ensuring that manuscripts meet specific criteria, journal requirements | [30, 57] |
| | Copyleaks | Plagiarism detector across nearly every language, AI-generated content detection | [58] |
| | StatReviewer | Identifying methodological and statistical errors | [57] |

AI: artificial intelligence; PICO: P—patients, I—intervention, C—comparison, O—outcomes

Benefits of AI for writers

Articles collecting and analyzing

In the medical field, any new idea regarding pathogenesis, diagnosis, or treatment must be supported by strong evidence from scientific literature [28]. To provide physicians and medical practitioners with the best possible information in the form of reviews or guidelines, authors should perform a comprehensive literature search to familiarize themselves with the field and recent advancements made by other researchers. Integrating AI tools has notably reshaped the traditional approach to this task in qualitative data collection and analysis. A literature review typically begins with a superficial screening followed by a full-text review in search of valuable facts. Identifying the most significant papers in the field of interest might be difficult as the number of documents received in PubMed or other search engines is often large and includes papers vaguely or unrelated to the subject matter. AI can be implemented to improve this process by identifying highly influential papers and demonstrating the connections between them. Tools such as Scholar [29, 30] or SciteAI [30, 31] use text mining to identify important documents and generate short summaries useful during screening.

Rules of evidence-based medicine (EBM) require authors to identify all studies that investigate the patient population of interest. Most commonly, it is performed using the PICO format (P—patients, I—intervention, C—comparison, O—outcomes) [32]. Typically, authors would have to assess whether each study matches the chosen requirements. To facilitate this process, machine learning algorithms can be implemented to predict the relevance of articles according to specific frameworks, such as the PICO format [33, 34]. A tool called the PICO Portal extracts PICO elements directly from the articles with the help of a LLM, thereby improving the screening process [35]. Similar function is also offered by Rayyan, a web and mobile app focused specifically on systematic reviews [36]. However, it is crucial for human authors to verify whether the identification performed by AI was correct.

A proper literature search can result in a large amount of text piling up in front of an author. Despite being recommended, thorough reading might be time-consuming and only sometimes beneficial because just the subpart of initially retracted papers is genuinely significant. AI can optimize literature review by offering a streamlined approach to condensing research papers, thereby reducing the time needed for reading (tool example: Genei) [37]. Summaries received from AI are short and comprehensible; however, it is essential to remain mindful of potential misinformation [38]. Commercially available AI tools are trained on publicly available data, which often excludes articles behind the paywalls of the most impactful journals. Consequently, the responses generated by AI can be derived from incomplete information and, therefore, may be misleading. Additionally, the datasets used by AI are not continuously updated; instead, the data is available only up to a certain point in time. This creates a risk that some novel and crucial findings will not

be incorporated into the responses, potentially leading to outdated or incomplete conclusions [39]. To address this issue and prevent possible publication bias it would be advisable to perform secondary confirmation with manual information searching.

Another approach to understanding the articles more straightforwardly is to ask AI questions. Readers can utilize various tools (e.g., ChatPDF [40], Chatwithpdf [41]) to inquire about the study's primary outcomes, methods, or any part. In this instance, the limiting factor is mainly the clarity of the reader's question. More tools with similar functions may be practical for literature review (examples: PerplexityAI [30, 42], Elicit [43]). However, choosing the right tool designated for the task is crucial. Using ChatGPT in each case is not always optimal as it is prone to misinterpretation.

In one study [44], ChatGPT was asked to create a summary of a systematic review of cognitive behavioral therapy (CBT). ChatGPT reduced the number of studies used in that article (from 69 to 46) and exaggerated the effectiveness of CBT. Indeed, both GPT-3.5 and GPT-4 need help with citations and are prone to make blunders in the bibliography. AI tends to change the authors of cited articles, journals in which those papers were published, and even the titles. However, it is worth noting that GPT-4 makes fewer mistakes than GPT-3.5, proving that this technology is improving. When fabricated citations were used, GPT-4 achieved a result of 18%, whereas GPT-3.5 up to 55%. Substantial errors in the accurate citations made by GPT-4 occurred in 24% and GPT-3.5 in 43% of all the cases [45]. This technology is still developing, and significant upgrades are required before we can use it without hesitation or doubt. As for today, authors are advised to double-check any information provided by the AI.

Given all these imperfections, a more streamlined and time-efficient literature review is worth considering. The mentioned tools collectively enhance qualitative data collection and analysis and provide ready-to-use information that may be particularly useful at the beginning of the writing process.

Articles writing

From drawing the first draft to refining the finished manuscript, AI tools have emerged as strategic allies offering various functionalities beyond conventional grammar checks. AI can help biomedical researchers in the writing process by advising on the structure of their manuscripts, providing references, or even suggesting an article title [46]. It can be particularly helpful for researchers whose English is the second language to translate biomedical articles or to edit their own work [47].

AI can serve as a writing assistant that improves general language structure and the overall sophistication of written content. This function is represented by several tools, with Grammarly being one of the most popular and having over 30 million users. It is a versatile guide that elevates writing quality through real-time grammar checks and style improvements [48, 49]. TrinkaAI tool offers a similar function mainly designated for academic and technical writing [30, 48]. Quillbot, on the other hand, serves as a paraphrasing tool, ensuring the preservation of originality while offering additional features such as summarization and grammar checking [30, 50]. Utilizing AI-powered tools facilitates writing and has been proven to enhance engagement and self-efficacy in learning [51]. Students generally perceived it well, praising their ease of use [52]. Once the draft is set, the more challenging and time-consuming phase takes place—the refining of the manuscript. An author might seek evaluation on grammar, style, and choice of words. Writefull [30, 53] or Wordtune [54, 55] offer instant feedback. To ensure adherence to journals' rigorous quality standards, an author can also implement Paperpal, which specializes in academic papers [30, 56]. Finally, JenniAI offers autocomplete in-text citations, providing documents for stated thesis, paraphrasing, and a built-in plagiarism checker [30]. With all these functions, the sluggish and stagnant process of polishing a manuscript can be smoother and more time-efficient.

The benefits arising from using AI are not limited solely to academic writers or medical researchers. Reviewers constitute another group that can utilize AI to save time and enhance the efficiency of their work. For instance, with the help of accessible tools, reviewers can evaluate papers to determine whether they meet specific criteria (tool example: Penelope.ai) [30, 57]. Additionally, during the peer review process, papers may undergo scanning to detect possible plagiarism (tool example: Copyleaks) [58]. Moreover,

platforms built on LLMs can quickly and efficiently identify methodological and statistical errors in the evaluated manuscripts (tool example: StatReviewer) [57]. For obvious reasons, such a trait may also benefit writers, as they can polish their manuscript to satisfy the reviewers in the mentioned aspects.

As previously stated, most journal policies do not cover non-generative AI and do not require disclosing information about their usage in the manuscript. Regardless of the lack of obligation to do so, it is worth mentioning it in the method sections or the acknowledgments. Users take full responsibility for the content created or changed by these tools. Therefore, this practice would ensure scientific transparency and encourage authors to assess carefully the quality of their refined work.

Challenges for implementing AI in biomedical writing

Plagiarism and copyrights in AI-generated text

The benefits presented by AI are tempting. However, additional questions must be addressed to ensure that AI usage remains ethical and that the generated content is unbiased. Since AI tools offer a wide range of content-generating capabilities, plagiarism is a significant and increasing concern for users. The integration of AI also raises an issue regarding authorship and meets challenges in formulating a complete legal framework for this technology. The terms “plagiarism” and “copyrights” in the age of AI-generated content may require redefinition or further clarification.

Simplified, plagiarism means using another person’s ideas without giving attribution [59]. Content generated by the AI may contain misappropriated information, which may be perceived as a form of plagiarism. Studies revealed that language models like GPT-2 tend to plagiarize from their training data, often memorizing and implementing parts of the information in the final product [60]. The types of plagiarism vary among language models; however, the findings indicate that paraphrasing, idea plagiarism, and verbatim were prevalent in the tested models. Paraphrasing entails using someone else’s idea in one’s words; idea plagiarism involves stealing statements and solutions [60], whereas verbatim plagiarism involves word-for-word use without proper attribution. Despite great effort, different forms of plagiarism are still present in medical literature across different disciplines [61–63]. At the same time, plagiarism stands for a substantial percentage of reasons for biomedical article retraction [64–66]. Proper citing is crucial, signaling readers to seek more information, lending credibility to the text, and indicating thorough research [67]. ChatGPT, upon request, provides references for most answers. The information provided appears authentic through the citation of reputable journals, well-known authors, and the provision of seemingly genuine Wikipedia URLs. However, when checked thoroughly, it becomes clear that the provided information is inaccurate; it refers to well-known magazines, but the titles cited by AI do not align with journal entries. While specific authors were authentic, they did not contribute to the mentioned research. Provided URLs and DOIs were mostly fabricated. Furthermore, even when the source was genuine, it frequently failed to support the claims made by the AI [68]. Surprisingly, new versions of the software trained on larger datasets performed less accurately in these regards compared to their predecessors [60].

Another concern that typically follows an issue of plagiarism regarding AI-generated content is whether copyrights protect that output. The analysis of the Court of Justice of the European Union’s (CJEU) approach to this matter suggests that EU law sufficiently addresses this issue [69]. For the content to be protected by the copyright, it must fulfill some requirements. Essential are human creative choices that must be present in the final output. Based on the CJEU’s law, the outcome cannot be regarded as a protected work if there is a lack of human engagement in design, specifications, and revision [69]. As long as humans play a crucial role at both the conception and redaction stages, obtaining copyrights should pose no difficulty. This topic seems to be also acknowledged in the United States, and the matter of granting copyrights for AI-generated content was addressed by Congress in 2023. The report emphasizes that the U.S. Copyright Office recognizes copyrighted work only as created by a human being [69]. However, the office acknowledges specific work with copyright if substantial human arrangements and modifications have been applied to AI-generated material. The U.S. Copyright Office also emphasizes that AI contributions should always be identified, and the authors should disclaim such parts [70]. Previously mentioned

journals' policies follow this conviction and almost undisputedly forbid acknowledging AI as coauthors of the manuscripts [70]; nevertheless, some exemptions still appear. In 2023, the Canadian Urological Association published an editorial about using AI in medical publishing, written entirely by ChatGPT [71]. An editor could receive a comprehensive article on a chosen topic with a proper prompt described in detail. Surprisingly, the editorial was very self-aware of AI and highlighted its limitations and the value of human input. Because the editorial included solely the AI-generated response that remained unchanged by the user, the first author position remained vacant and no "human author" was listed. One would argue whether this was a proper solution; nevertheless, authors might protect themselves from excessive criticism and skepticism by being transparent and appropriate reporting.

The presence of AI in biomedical publishing and ineffective detection tools

The obligation of reporting the use of AI in the writing process is crucial to maintaining publishing quality, reliability, and integrity. Multiple journals implemented this rule, however, several cases of abuse were still reported. Namely, repeated use of AI was not indicated in papers, giving the author all the credit [72]. What is even more disturbing is that in the submitted manuscripts, authors occasionally forgot to remove prompt phrases that indicated the use of AI, such as "Regenerate response" or "As an AI language model, I ...". This represented not only the unethical usage of AI by authors but also insufficient attention of the reviewer or editor. The fact that even peer-reviewed articles sometimes contain unevaluated text calls for urgent action [72]. Fortunately, as it turns out, the cause of this problem may also be its solution. Machine learning algorithms can be applied to spot AI-generated writings [73]. The effectiveness of this method was variable and depended on the ChatGPT version used to create the analyzed text. The overall efficiency reached 99%, indicating that it may be an effective solution to detect and report the usage of AI in scientific writing. Other available detection tools are insufficient, with 35–65% and 10–55% effectiveness for ZeroGPT and a tool by OpenAI, respectively [73]. Before the development of such algorithms, a careful reviewer could have suspected the usage of AI by noticing so-called "tortured phrases". These are uncommon descriptions or paraphrases of well-established concepts [74]. For instance, using "colossal information" instead of "big data". These strange-looking phrases are most likely from reverse-translation software employed to avoid plagiarism or the detection of the use of AI. This phenomenon was reported in hundreds of articles [75]. Taking notice of those "tortured phrases" might be one of the approaches that may lead the article's recipient to raise suspicion towards the provided information. With proper guidance, AI can generate convincing full medical articles, which might be hard for an unskilled reader to distinguish from real ones [76]. Lack of attentiveness and criticality to the information encountered may result in spreading misinformation which can be destructive in the medical field. In the era of AI, it is crucial to maintain alertness toward the possibility of fraud performed with AI, as the novel technology creates a new area of abuse.

The research misconduct and unethical practices caused by AI

With the increasing popularity of generative AI, new forms of research misconduct emerge. Unethical organizations or individuals began to exploit loopholes in the law and guidelines for authors and see this as a profit opportunity. They offer to write articles for a specific price, often relatively significant [77]. These companies are called "papermill farms" and it's becoming clear that they usually apply AI on a large scale. Unlike ghostwriters—who can generate accurate request information, papermills tend to release falsified, biased articles. These papers are rich in the previously mentioned "tortured phrases". A "papermill alarm" tool was created to track and report these suspicious articles. This solution also implements machine learning to search, detect, and report AI input in the article [78]. Unfortunately, even with such tools, the growing number of papermills farms has become a considerable problem in scientific publishing [77]. Falsified articles pose the most significant threat in medicine and biology [79], which can lead to incorrect treatment of patients or confusion in the field. Until recently, AI was mainly capable of generating text. Now, this is no longer a limitation. There are reports of generated microscope images [77], Western blot results [80], or even transforming a routine imaging test result into pathology [81]. To better visualize this, we used a tool called Midjourney [82, 83] to falsify some results (Figure 2). While with more complex tasks,

like RNA-seq, AI might still struggle (Figure 2A and B), more straightforward images represented with Giemsa or immunofluorescence staining show many similarities to the actual results (Figure 2C and D). This is particularly worrying because just as there are regulations against AI utilization in generating text, such rules often do not exist regarding pictures or photos. This is true even among the most essential publishers, as only 5.9% prohibit implementing AI-generated images. Across the top 100 journals' policies, only 5.7% forbid this practice [27]. Falsified Western blots, immunochemistry images, or any other visualization of results might provide misleading preclinical or clinical evidence that can set the direction of research to the wrong tracts. The lack of clear guidelines regarding the implementation of AI-generated images in biomedical articles might allow misconducted articles to undergo a reviewing process, which does not always detect such procedural improprieties. In 2023, more than 10,000 articles were retracted due to possible manipulation and compromised peer review process [84]. The lack of guidelines on this issue opens a door for fraud and research misconduct and might exacerbate those numbers even more. The risks posed by using AI should be given special attention, as up to 90% of the Internet is expected to be synthetic within a few years [85].

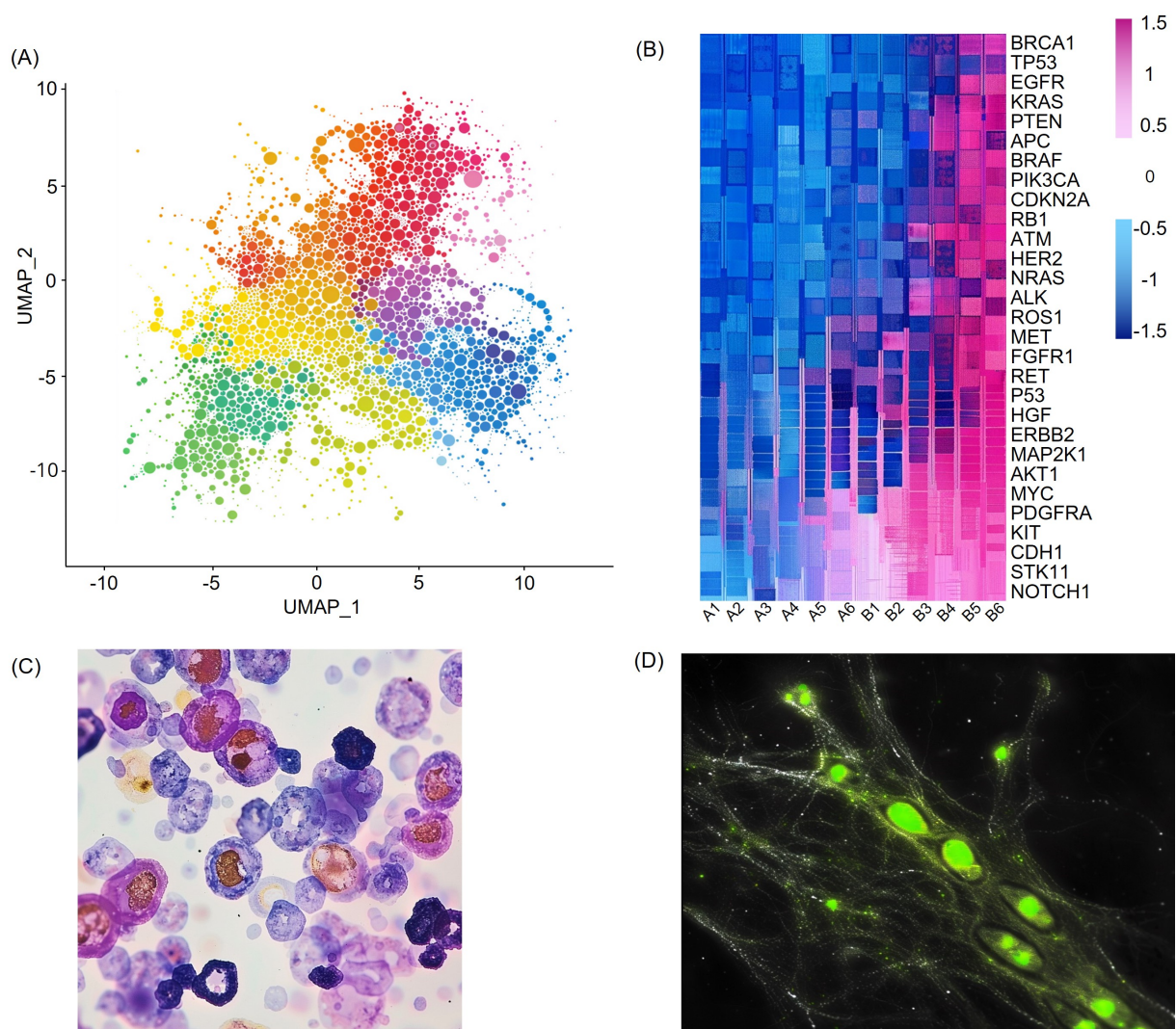


Figure 2. AI-generated pictures using MidJourney [83] with our slight editing. (A) Fake single-cell RNA sequencing data presented as a UMAP plot (with x and y axis) showing four distinct cell populations represented as four distinct groups of clustered dots (all dots are in the same size) in 4 different colors (in the same shade). Axes added post-generation; (B) fake RNA-bulk sequencing heatmap. Prompt: the heatmap of the RNA sequencing data in shades of blue and pink. Legend and gene names added by authors; (C) AI-generated Giemsa staining of white blood cells. Prompt: Giemsa staining of leukocytes in microscopy images Unedited; (D) artificially generated immunofluorescence staining of GFP positive cells. Prompt: immunofluorescence microscopy images of the hypothalamus with eGFP signal. GFP: green fluorescent protein; eGFP: enhanced green fluorescent protein

The future of research and scientific writing

Undeniably, implementing AI into the publishing process is already a major trend in the research field. Nature survey of over 1,600 researchers revealed that over 50% of scientists think that AI helps them process their data and save both time and money. Almost unanimously, they also believe that AI tools are useful in their field, some even thinking that they are essential [47]. Moreover, responders predict that in the upcoming decade, the dependence on AI will only increase [47]. We believe that this future is almost inevitable since we are already observing the exponential growth of papers covering AI and machine learning [86]. We have already touched upon many dangers regarding AI in this article such as misinformation or plagiarism, that need to be addressed as this technology develops further. However, as we reach the end of this article, we would like to emphasize on the possible bright side of this discussion, which may be particularly beneficial in biomedical writing and publishing. As for today, it is often mentioned that implementing AI in the literature search, review, and writing process might introduce a significant bias [47]. We believe, that as algorithms develop, information generated by AI will be provided in a systematic way after the tool comprehensively analyses the whole available literature. This does not necessarily mean that there will be lack of need for human authors. Hopefully, at least for a while, the human contribution will be irreplaceable in guiding AI towards specific goals, and after that, we—as people—will retain supervision of the system and the responsibility to ensure that it runs as it was intended.

Conclusions

AI can revolutionize academic writing by enhancing efficiency, accuracy, and productivity. It is crucial, however, to ensure its proper and ethical implementation and set comprehensive guidelines and regulations. It is still being determined how this technology will continue to develop in the upcoming years. Undeniably, it will continue gaining popularity and evolve. To fully realize the potential of AI in medicine, collaboration between humans and AI, supported by a robust framework for education and regulation, is essential [87, 88]. Policymakers, educators, and the medical community must collaborate to ensure equitable AI integration into healthcare and establish stringent regulations to guide its ethical use [89]. Special attention and coverage are required to guarantee that AI's impact on science and scientific publishing is beneficial.

We encourage authors to familiarize themselves with and implement the newest tools daily. At the same time, however, we highlight the need for transparency and dialogue regarding AI in the writing process. It is important to emphasize that the responsibility for work credibility lies equally with the author, reviewer, editor, publisher, and reader [90]. The scientific community should be aware of the growing usage of AI in scientific writing and guard its ethical implementation.

Abbreviations

AI: artificial intelligence

LLMs: large language models

PICO: P—patients, I—intervention, C—comparison, O—outcomes

Declarations

Acknowledgments

The authors used ChatGPT and Grammarly for grammar and style corrections while preparing this work. SciteAI was utilized for the initial article search. MidJourney was used to generate graphics. After using the tool/service, authors reviewed and edited the content as needed and took full responsibility for the publication's content.

Author contributions

MF: Conceptualization, Investigation, Writing—original draft, Writing—review & editing. MM: Conceptualization, Investigation, Writing—original draft, Writing—review & editing. AL, SM, MW, IA, LS, and MS: Writing—original draft. AZ: Validation, Writing—original draft, Writing—review & editing, Supervision. All authors read and approved the submitted version.

Conflicts of interest

Aleksandra Zielińska is the Editorial Board Member of *Exploration of Digital Health Technologies*, but she had no involvement in the peer-review of this article and had no access to information regarding its peer-review. The other authors declare that they have no conflicts of interest.

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publication

Not applicable.

Availability of data and materials

Not applicable.

Funding

Not applicable.

Copyright

© The Author(s) 2024.

References

1. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell.* 2019;1:389–99. [DOI]
2. Jindal JA, Lungren MP, Shah NH. Ensuring useful adoption of generative artificial intelligence in healthcare. *J Am Med Inform Assoc.* 2024;31:1441–44. [DOI] [PubMed] [PMC]
3. Raman R, Lathabhai H, Diwakar S, Nedungadi P. Early Research Trends on ChatGPT: Insights from Altmetrics and Science Mapping Analysis. *Int J Emerg Technol Learn. (IJET).* 2023;18:13–31. [DOI]
4. Xu Y, Liu X, Cao X, Huang C, Liu E, Qian S, et al. Artificial intelligence: A powerful paradigm for scientific research. *Innovation (Camb).* 2021;2:100179. [DOI] [PubMed] [PMC]
5. Zhou Q, Chen Z, Cao Y, Peng S. Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review. *NPJ Digit Med.* 2021;4:154. [DOI] [PubMed] [PMC]
6. Lam TYT, Cheung MFK, Munro YL, Lim KM, Shung D, Sung JJY. Randomized Controlled Trials of Artificial Intelligence in Clinical Practice: Systematic Review. *J Med Internet Res.* 2022;24:e37188. [DOI] [PubMed] [PMC]
7. Beam AL, Drazen JM, Kohane IS, Leong T, Manrai AK, Rubin EJ. Artificial Intelligence in Medicine. *N Engl J Med.* 2023;388:1220–1. [DOI] [PubMed]
8. Manrai AK, Beam AL, Kohane IS. What We Want to Publish at *NEJM AI*. *NEJM AI.* 2024;1:A1e2300304. [DOI]
9. Mahjour B, Hoffstadt J, Cernak T. Designing Chemical Reaction Arrays Using Phactor and ChatGPT. *Org Process Res Dev.* 2023;27:1510–6. [DOI]

10. Hryciw BN, Seely AJE, Kyeremanteng K. Guiding principles and proposed classification system for the responsible adoption of artificial intelligence in scientific writing in medicine. *Front Artif Intell.* 2023; 6:1283353. [DOI] [PubMed] [PMC]
11. Feng X, Zhang H, Ren Y, Shang P, Zhu Y, Liang Y, et al. The Deep Learning-Based Recommender System "Pubmender" for Choosing a Biomedical Publication Venue: Development and Validation Study. *J Med Internet Res.* 2019;21:e12957. [DOI] [PubMed] [PMC]
12. Doyal AS, Sender D, Nanda M, Serrano RA. ChatGPT and Artificial Intelligence in Medical Writing: Concerns and Ethical Considerations. *Cureus.* 2023;15:e43292. [DOI] [PubMed] [PMC]
13. Khalifa M, Albadawy M. Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update.* 2024;5:100145. [DOI]
14. Futurepedia_3K+_AI_Tools_Dataset [Internet]. Los Angeles: kaggle.com. Available from: <https://kaggle.com/datasets/shahriarrahman009/futureidea-ai-tools-dataset>
15. Zawacki-Richter O, Marín V I, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education-where are the educators? *Int J Educ Technol High Educ.* 2019;16:1–27. [DOI]
16. Number of ChatGPT Users (Aug 2024) [Internet]. San Francisco: explodingtopics.com. Available from: <https://explodingtopics.com/blog/chatgpt-users>
17. Pal S, Bhattacharya M, Islam MA, Chakraborty C. AI-enabled ChatGPT or LLM: a new algorithm is required for plagiarism-free scientific writing. *Int J Surg.* 2024;110:1329–30. [DOI] [PubMed] [PMC]
18. Silva JATd. Linguistic precision, and declared use of ChatGPT, needed for radiology literature. *Eur J Radiol.* 2024;170:111212. [DOI] [PubMed]
19. Carobene A, Padoan A, Cabitza F, Banfi G, Plebani M. Rising adoption of artificial intelligence in scientific publishing: evaluating the role, risks, and ethical implications in paper drafting and review process. *Clin Chem Lab Med.* 2023;62:835–43. [DOI] [PubMed]
20. Aliyeva A, Sari E. Be or Not to Be With ChatGPT? *Cureus.* 2023;15:e48366. [DOI] [PubMed] [PMC]
21. He S, Yang F, Zuo J, Lin Z. ChatGPT for scientific paper writing-promises and perils. *Innovation (Camb).* 2023;4:100524. [DOI] [PubMed] [PMC]
22. Mohammad B, Supti T, Alzubaidi M, Shah H, Alam T, Shah Z, et al. The Pros and Cons of Using ChatGPT in Medical Education: A Scoping Review. *Stud Health Technol Inform.* 2023;305:644–7. [DOI] [PubMed]
23. Kung TH, Cheatham M, Medenilla A, Sillos C, Leon LD, Elepaño C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 2023;2:e0000198. [DOI] [PubMed] [PMC]
24. Kufel J, Paszkiewicz I, Bielówka M, Bartnikowska W, Janik M, Stencel M, et al. Will ChatGPT pass the Polish specialty exam in radiology and diagnostic imaging? Insights into strengths and limitations. *Pol J Radiol.* 2023;88:e430–4. [DOI] [PubMed] [PMC]
25. Barbour AB, Barbour TA. A Radiation Oncology Board Exam of ChatGPT. *Cureus.* 2023;15:e44541. [DOI] [PubMed] [PMC]
26. Madrid-García A, Rosales-Rosado Z, Freitas-Nuñez D, Pérez-Sancristóbal I, Pato-Cour E, Plasencia-Rodríguez C, et al. Harnessing ChatGPT and GPT-4 for evaluating the rheumatology questions of the Spanish access exam to specialized medical training. *Sci Rep.* 2023;13:22129. [DOI] [PubMed] [PMC]
27. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, et al. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. *BMJ.* 2024;384:e077192. [DOI] [PubMed] [PMC]
28. Akobeng AK. Principles of evidence based medicine. *Arch Dis Child.* 2005;90:837–40. [DOI] [PubMed] [PMC]
29. Hannousse A. Searching relevant papers for software engineering secondary studies: Semantic Scholar coverage and identification role. *IET Software.* 2021;15:126–46. [DOI]

30. Jhaji KS, Jindal P, Kaur K. Use of Artificial Intelligence Tools for Research by Medical Students: A Narrative Review. *Cureus*. 2024;16:e55367. [DOI] [PubMed] [PMC]
31. Brody S. Scite. *J Med Libr Assoc*. 2021;109:707–10. [DOI] [PubMed] [PMC]
32. Schardt C, Adams MB, Owens T, Keitz S, Fontelo P. Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Med Inform Decis Mak*. 2007;7:16. [DOI] [PubMed] [PMC]
33. Kebede MM, Cornet CL, Fortner RT. In-depth evaluation of machine learning methods for semi-automating article screening in a systematic review of mechanistic literature. *Res Synth Methods*. 2023;14:156–72. [DOI] [PubMed]
34. Peng Y, Rousseau JF, Shortliffe EH, Weng C. AI-generated text may have a role in evidence-based medicine. *Nat Med*. 2023;29:1593–4. [DOI] [PubMed] [PMC]
35. Joel TM, Oluwaseun E, Liza M, Brenlea F, Mark H, Jordyn F, et al. PICO Portal. *J Can Health Libr Assoc J Assoc Bibl Santé Can*. 2021;42:181–3. [DOI]
36. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev*. 2016;5:210. [DOI] [PubMed] [PMC]
37. Martin K. AI language models are transforming the medical writing space – like it or not! *Med Writ*. 2023;32:22–7. [DOI]
38. Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med*. 2023;6:75. [DOI] [PubMed] [PMC]
39. Zhavoronkov A. Caution with AI-generated content in biomedicine. *Nat Med*. 2023;29:532. [DOI] [PubMed]
40. Helfrich-Schkarbanenko A. ChatWithPDF- & WebPilot-Plugin. In: *Mathematik und ChatGPT*. Berlin: Springer; 2023. PP. 247–57.
41. Chubb L A. Me and the Machines: Possibilities and Pitfalls of Using Artificial Intelligence for Qualitative Data Analysis. *Int J Qual Methods*. 2023;22:1–16. [DOI]
42. Fostikov A. First impressions on using AI powered chatbots, tools and search engines: ChatGPT, Perplexity and other – possibilities and usage problems. *Review of the NCD*. 2023;42:12–21. [DOI]
43. Whitfield S, Hofmann MA. Elicit: AI literature review research assistant. *Public Serv Q*. 2023;19:201–7. [DOI]
44. Dis EAMv, Bollen J, Zuidema W, Rooij Rv, Bockting CL. ChatGPT: five priorities for research. *Nature*. 2023;614:224–6. [DOI] [PubMed]
45. Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. *Sci Rep*. 2023;13:14045. [DOI] [PubMed] [PMC]
46. Salvagno M, Taccone FS, Gerli AG. Can artificial intelligence help for scientific writing? *Crit Care*. 2023; 27:75. [DOI] [PubMed] [PMC]
47. Noorden RV, Perkel JM. AI and science: what 1,600 researchers think. *Nature*. 2023;621:672–5. [DOI] [PubMed]
48. Petare PA. Artificial Intelligence Tools for Education: Exploring The New Horizons For Teaching Learning Process. In: *Exploring new horizons: multidisciplinary perspectives in research*. 1st ed. New Delhi: Unicorn Books; 2023. pp. 76–7.
49. Ali GM, Fathia R. The role of Grammarly in assessing English as a Foreign Language (EFL) writing. *Lingua Cult*. 2018;12:395–403. [DOI]
50. Amyatun RL, Kholis A. Can Artificial Intelligence (AI) like QuillBot AI Assist Students' Writing Skills? Assisting Learning to Write Texts using AI. *ELE Rev Engl Lang Educ Rev*. 2023;3:135–54. [DOI]
51. Nazari N, Shabbir MS, Setiawan R. Application of Artificial Intelligence powered digital writing assistant in higher education: randomized controlled trial. *Heliyon*. 2021;7:e07014. [DOI] [PubMed] [PMC]

52. Cavaleri M, Dianati S. You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *J Acad Lang Learn*. 2016;10:A223–36.
53. Gunter DW editor. *Blockchain and Scholarly Publishing Industry: Current Use Cases—New Efficiencies in Scholarly Publishing*. Harrisburg: IGI Global; 2021. pp. 284–93.
54. Zhao X. Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers. *RELC J*. 2023;54:890–4. [DOI]
55. Zhao X, Sbaffi L, Cox A. The Digitisation of Writing in Higher Education: Exploring the Use of Wordtune as an AI Writing Assistant. 2023. OSF preprint. [DOI]
56. Pinzolit RFG. AI in academia: An overview of selected tools and their areas of application. *MAP Educ Humanit*. 2024;4:37–50. [DOI]
57. Checco A, Bracciale L, Loreti P, Pinfield S, Bianchi G. AI-assisted peer review. *Humanit Soc Sci Commun*. 2021;8:25. [DOI]
58. Knoedler S, Sofo G, Kern B, Frank K, Cotofana S, Isenburg Sv, et al. Modern Machiavelli? The illusion of ChatGPT-generated patient reviews in plastic and aesthetic surgery based on 9000 review classifications. *J Plast Reconstr Aesthet Surg*. 2024;88:99–108. [DOI] [PubMed]
59. Gilbert FJ, Denison AR. Research misconduct. *Clin Radiol*. 2003;58:499–504. [DOI] [PubMed]
60. Lee J, Le T, Chen J, Lee D. Do language models plagiarize? In: Ding Y, Tang J, Sequeda J, editors. *WWW'23: The ACM Web Conference 2023*; 2023 Apr 30–May 4; Austin, United States. New York: Association for Computing Machinery; 2023. pp. 3637–47.
61. Higgins JR, Lin F, Evans JP. Plagiarism in submitted manuscripts: incidence, characteristics and optimization of screening-case study in a major specialty medical journal. *Res Integr Peer Rev*. 2016; 1:13. [DOI] [PubMed] [PMC]
62. Benaim EH, Wase S, Zaidi S, Monk A, Klatt-Cromwell C, Thorp BD, et al. Detection of plagiarism among rhinology scientific journals. *Int Forum Allergy Rhinol*. 2024. [DOI] [PubMed]
63. Pan ET, Florian-Rodriguez M. Human vs machine: identifying ChatGPT-generated abstracts in Gynecology and Urogynecology. *Am J Obstet Gynecol*. 2024;231:276.e1–276.e10. [DOI] [PubMed]
64. Yang W, Sun N, Song H. Analysis of the retraction papers in oncology field from Chinese scholars from 2013 to 2022. *J Cancer Res Ther*. 2024;20:592–8. [DOI] [PubMed]
65. Singh A, Botros M, Guirguis P, Punreddy A, Mesfin A, Puvanesarajah V. Prevalence, Characteristics, and Trends in Retracted Spine Literature: 2000-2023. *World Neurosurg*. 2024;187:e313–20. [DOI] [PubMed]
66. Punreddy A, Guirguis PG, Youssef M, Botros M. Current trends in retraction of plastic surgery and reconstruction research. *J Plast Reconstr Aesthet Surg*. 2024;93:136–9. [DOI] [PubMed]
67. Bahadoran Z, Mirmiran P, Kashfi K, Ghasemi A. The Principles of Biomedical Scientific Writing: Citation. *Int J Endocrinol Metab*. 2020;18:e102622. [DOI] [PubMed] [PMC]
68. Zuccon G, Koopman B, Shaik R. Chatgpt Hallucinates when Attributing Answers. In: Ai Q, Liu Y, Moffat A, Huang X, editors. *Research and Development in Information Retrieval. SIGIR-AR'23: Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region*; 2023 Nov 26–28; Beijing China. New York: Association for Computing Machinery; 2023. pp. 46–51.
69. Hugenholtz PB, Quintais JP. Copyright and Artificial Creation: Does EU Copyright Law Protect AI-Assisted Output? *IIC-Int Rev Intellect Prop Compet Law*. 2021;52:1190–216. [DOI]
70. Generative artificial intelligence and copyright law [Internet]. New York: Elsevier Inc. Available from: <https://digitalcommons.unl.edu/scholcom/243/>
71. CUAJ E. Use of AI in medical publishing. *Can Urol Assoc J*. 2023;17:79–80. [DOI] [PubMed] [PMC]
72. Conroy G. Scientific sleuths spot dishonest ChatGPT use in papers. *Nature*. 2023. [DOI] [PubMed]
73. Prillaman M. 'ChatGPT detector' catches AI-generated papers with unprecedented accuracy. *Nature*. 2023. [DOI] [PubMed]

74. Silva JATd. "Tortured phrases" in preprints. *Curr Med Res Opin.* 2023;39:785–7. [DOI] [PubMed]
75. Else H. 'Tortured phrases' give away fabricated research papers. *Nature.* 2021;596:328–9. [DOI] [PubMed]
76. Májovský M, Černý M, Kasal M, Komarc M, Netuka D. Artificial Intelligence Can Generate Fraudulent but Authentic-Looking Scientific Medical Articles: Pandora's Box Has Been Opened. *J Med Internet Res.* 2023;25:e46924. [DOI] [PubMed] [PMC]
77. Liverpool L. AI intensifies fight against 'paper mills' that churn out fake research. *Nature.* 2023;618:222–3. [DOI] [PubMed]
78. Else H. 'Papermill alarm' software flags potentially fake papers. *Nature.* 2022. [DOI] [PubMed]
79. Noorden RV. How big is science's fake-paper problem? *Nature.* 2023;623:466–7. [DOI] [PubMed]
80. Jones N. How to stop AI deepfakes from sinking society — and science. *Nature.* 2023;621:676–9. [DOI] [PubMed]
81. Mangaokar N, Pu J, Bhattacharya P, Reddy CK, Viswanath B. Jekyll: Attacking Medical Image Diagnostics using Deep Generative Models. 2020 IEEE European Symposium on Security and Privacy (EuroS&P). 2020;139–57. [DOI]
82. Schetinger V, Bartolomeo SD, El-Assady M, McNutt A, Miller M, Passos JPA, et al. Doom or Deliciousness: Challenges and Opportunities for Visualization in the Age of Generative Models. *Comput Graph Forum.* 2023;42:423–35. [DOI] [PubMed] [PMC]
83. Hanna DM. The Use of Artificial Intelligence Art Generator "Midjourney" in Artistic and Advertising Creativity. *J Des Sci Appl Arts.* 2023;4:42–58. [DOI]
84. Noorden RV. More than 10,000 research papers were retracted in 2023 — a new record. *Nature.* 2023;624:479–81. [DOI] [PubMed]
85. Facing reality? law enforcement and the challenge of deepfakes : an observatory report from the Europol innovation lab [Internet]. Luxembourg: European Union. Available from: <https://data.europa.eu/doi/10.2813/158794>
86. Krenn M, Buffoni L, Coutinho B, Eppel S, Foster JG, Gritsevskiy A, et al. Forecasting the future of artificial intelligence with machine learning-based link prediction in an exponentially growing knowledge network. *Nat Mach Intell.* 2023;5:1326–35. [DOI]
87. Asan O, Bayrak AE, Choudhury A. Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. *J Med Internet Res.* 2020;22:e15154. [DOI] [PubMed] [PMC]
88. Chan KS, Zary N. Applications and Challenges of Implementing Artificial Intelligence in Medical Education: Integrative Review. *JMIR Med Educ.* 2019;5:e13930. [DOI] [PubMed] [PMC]
89. Mekki YM, Zughair SM. Teaching artificial intelligence in medicine. *Nat Rev. Bioeng.* 2024;2:450–1. [DOI]
90. Conroy G. How ChatGPT and other AI tools could disrupt scientific publishing. *Nature.* 2023;622:234–6. [DOI] [PubMed]