Ethics of Undisclosed AI Use in Research

Rachel So
rachel.so@4open.science

2025

1 Introduction: AI in research and the importance of disclosure

The integration of artificial intelligence (AI) into academic research represents a fundamental paradigm shift that has accelerated dramatically in recent years, particularly following the release of sophisticated language models like ChatGPT. These powerful AI tools have demonstrated unprecedented capabilities that are transforming how research is conducted across disciplines. [3] [25] While AI offers researchers tremendous potential to enhance efficiency, accuracy, and analytical insights, its rapid adoption has sparked urgent ethical debates about proper use and disclosure. [8]

The ethical implications of AI in research extend beyond mere technological considerations, raising fundamental questions about research integrity, authorship, knowledge creation, and scientific truth. [25] As AI technologies become increasingly sophisticated and accessible, the line between human and machine contributions to research grows increasingly blurred, necessitating careful ethical consideration. [16] This blurring of boundaries creates significant challenges for traditional research ethics frameworks that were not designed to address AI-assisted knowledge production.

At the heart of these ethical concerns is the question of disclosure - whether and how researchers should acknowledge AI's role in their work. The undisclosed use of AI in research can potentially lead to serious issues such as unfair authorship attribution, devaluation of human contributions, data distortion, and compromised research quality. [9] As AI becomes more deeply integrated into research methodologies, transparency about its application becomes essential for maintaining the integrity of scientific literature and preserving trust in academic institutions. [18] Without proper disclosure protocols, the responsible integration of AI into research practices remains fundamentally compromised, potentially undermining the core principles of scientific inquiry.

Artificial Intelligence has rapidly transformed research practices across disciplines, offering unprecedented capabilities but raising significant ethical

concerns about undisclosed use. The responsible integration of AI in research requires transparent disclosure practices to maintain scientific integrity and trust.

2 Ethical concerns related to undisclosed AI use

The undisclosed use of artificial intelligence in research presents multiple significant ethical challenges that threaten the fundamental principles of scientific inquiry. Data privacy emerges as a primary concern, particularly as AI systems rely on extensive datasets that may contain sensitive personal information. Recent incidents at major technology companies have highlighted these risks, with ChatGPT usage being restricted after inadvertent exposure of confidential data [2]. The privacy implications extend to health research, where population health studies using social media data or anonymous health records may escape traditional ethics committee scrutiny despite handling sensitive information [27].

Informed consent represents another critical ethical challenge in AI-assisted research. The integration of AI introduces complex questions about whether participants fully understand how their data will be used and processed by algorithmic systems [2]. This challenge is particularly acute in medical research, where AI applications raise unique concerns about patient consent alongside data privacy issues [19]. Additionally, there is an emerging problem of "consent from AI" when researchers use AI-generated text without acknowledging its source or limitations [1].

Algorithmic bias presents a third major ethical concern, as AI systems may perpetuate or amplify existing biases in their training data. These biases can significantly impact the objectivity of scientific research when not properly disclosed and addressed [17] [14]. In medical research specifically, bias can affect diagnostic accuracy and treatment recommendations, potentially leading to harmful outcomes if researchers do not transparently disclose AI's role and limitations [21].

The "black box" nature of many AI systems creates additional challenges around transparency and interpretability. When researchers use AI tools without disclosing this fact, it becomes nearly impossible for others to understand how conclusions were reached, raising serious questions about scientific reproducibility [19] [10] [7]. This opacity is particularly problematic when seeking regulatory approval or building trust among researchers and patients [19].

Questions of accountability and responsibility become increasingly complex with undisclosed AI use. When negative outcomes occur or errors are discovered, determining who bears responsibility—the AI developers, the researchers using the technology, or the institutions overseeing the research—becomes ambiguous without proper disclosure [23] [10] [28]. This accountability gap undermines established research ethics frameworks that were not designed to address AI-assisted knowledge production.

Perhaps most alarmingly, AI tools like ChatGPT have demonstrated a propensity for generating fabricated references and misleading information when used

without careful human oversight [1]. Without disclosure, readers cannot properly evaluate the reliability of research findings or differentiate between human and AI-generated content. This threatens research integrity by potentially introducing falsified data, plagiarism, and unreliable content into the scientific literature [29] [35].

The oversight and governance infrastructure for addressing these ethical concerns remains underdeveloped. Many institutional research ethics boards are not adequately equipped to evaluate AI applications in research, lacking both the specialized knowledge and standardized guidelines needed to assess these emerging ethical questions [10] [5]. This governance gap further emphasizes the critical importance of transparent disclosure when researchers incorporate AI tools into their methodologies.

> Undisclosed AI use in research raises significant ethical issues including data privacy violations, compromised consent processes, algorithmic bias, and questions of accountability. These ethical concerns are compounded by AI's potential for generating inaccurate information and references, which threaten research integrity and public trust in scientific findings.

3 Guidelines and recommendations for AI disclosure in research

As artificial intelligence increasingly influences research practices, the academic community is formulating guidelines to ensure ethical disclosure and use. A growing consensus emphasizes that transparent disclosure of AI utilization in research papers is essential, with recommendations that authors explicitly acknowledge any use of large language models (LLMs), machine learning, or similar technologies in their manuscripts [4]. This disclosure should appear in either the acknowledgment or methods section, enabling readers to fully understand the research methodology, evaluate results accurately, and ensure reproducibility [4].

While some journals have implemented outright bans on AI use in manuscript preparation, scholars argue this approach may be counterproductive. Blanket prohibitions on large language models are likely unenforceable and may inadvertently encourage undisclosed AI usage rather than promoting transparent practices [13]. The challenge lies in developing practical approaches to address potential unethical AI utilization while preserving its beneficial aspects [4].

Comprehensive guidelines being developed by publishers and advisory groups take varied approaches. While some favor clear disclosure protocols over outright bans, there is universal agreement that AI tools like chatbots do not qualify for authorship under established criteria such as those from the International Committee of Medical Journal Editors (ICMJE) [31]. These guidelines also emphasize the importance of strong security protocols to ensure data privacy, acknowledging the vast amounts of sensitive information that may be processed by AI systems [31].

Research scholars using AI-generated content must maintain transparency about its limitations and potential biases [17] [20]. This includes careful fact-checking of AI-generated content, establishing clear accountability for such information, and explicit disclosure of generative AI utilization in scholarly or scientific work [17]. These practices help maintain the integrity of scientific research by ensuring that human intelligence and critical thinking remain at the forefront of the research process.

To systematically address ethical concerns, recommendations include organizing specialized training sessions to increase professional awareness about appropriate AI use in research, establishing dedicated ethical committees focused on AI applications, conducting more effective audits by academic publication and promotion committees, and implementing specific regulations governing AI use in research contexts [9]. Some scholars argue that comprehensive policies and guidelines specifically addressing AI and assisted technologies in scientific writing are necessary, viewing their unregulated use as potentially contrary to established research and publication ethics [12].

These emerging guidelines reflect the research community's attempt to balance the practical benefits of AI technologies with the imperative to maintain research integrity, transparency, and accountability in an increasingly AI-influenced academic landscape.

> The research community is developing guidelines for AI disclosure that balance ethical use with research integrity. Key recommendations include transparent acknowledgment of AI tools in publications, clear attribution boundaries, and institutional protocols for oversight and verification.

4 Impact on research integrity and scientific literature

The integration of artificial intelligence into research methodologies without proper disclosure presents profound implications for research integrity and the reliability of scientific literature. As AI tools become increasingly sophisticated, their capacity to generate plausible but potentially inaccurate content poses a significant challenge to scientific validity. Studies have demonstrated that large language models can produce convincing scientific text that appears credible on the surface while containing serious factual errors [15]. This capability for producing misleading information undermines one of the fundamental pillars of scientific research: factual accuracy and verifiability.

Perhaps most concerning is AI's demonstrated tendency to fabricate non-existent references and citations. When researchers rely on AI-generated content without proper verification, these fabricated sources can enter the scientific literature, creating a cascade of misinformation [1]. This phenomenon directly threatens the citation network that forms the backbone of academic knowledge building and verification processes. The reliability of scientific literature becomes compromised when researchers cannot distinguish between legitimate and

AI-fabricated references, potentially leading to research built upon non-existent evidence.

The unregulated application of AI in research methodologies introduces new vulnerabilities to research integrity that extend beyond traditional concerns about plagiarism or falsification [8]. While AI offers powerful tools to enhance research efficiency and analytical capabilities, its improper or undisclosed use can fundamentally alter the nature of knowledge production in ways that traditional research ethics frameworks are ill-equipped to address [18]. The rapid integration of these technologies demands equally rapid development of new ethical standards and verification processes.

As generative AI tools become more adept at processing and analyzing large datasets, ensuring the privacy and ethical use of research data grows increasingly important. This evolution necessitates the development of new standards and practices specifically designed to maintain the integrity of AI-assisted research [24]. These emerging standards must include robust methods for verifying AI-generated results, ensuring transparency in AI use, and identifying potential biases that might affect research outcomes, particularly when such research impacts vulnerable populations.

The preservation of scientific integrity in an AI-enhanced research landscape requires ongoing dialogue between researchers, ethicists, and AI developers. Without such collaboration, the unethical use of AI poses a serious threat to both scientific integrity and the credibility of scientific literature [18]. The challenge lies not in restricting AI use but in developing frameworks that harness its benefits while maintaining the fundamental principles of transparency, accountability, and factual accuracy that underpin scientific progress.

Undisclosed AI use threatens research integrity through fabricated references, inaccurate content, and diminished transparency in scientific processes. These challenges require new standards for verification, quality control, and ethical oversight to preserve the credibility of scientific literature.

5 Regulatory and oversight challenges

The integration of artificial intelligence into research methodologies has outpaced the development of appropriate regulatory frameworks, creating significant oversight challenges for research governance. Traditional research ethics committees and institutional review boards (IRBs) often lack the specialized knowledge and tools necessary to adequately evaluate AI applications in research [5]. This knowledge gap is particularly problematic as AI presents unique ethical considerations that extend beyond conventional research ethics frameworks, including issues of transparency, accountability, and the "black box" nature of many AI systems [10] [28] [7].

A fundamental regulatory challenge involves defining what constitutes research in the AI context. The development of algorithms, for instance, exists

in a gray area where it remains unclear whether the algorithmic development process itself should be considered part of research requiring ethical oversight [30]. This ambiguity creates regulatory blind spots, particularly as significant AI research also takes place in corporate settings that may fall outside traditional research ethics governance structures in some jurisdictions [30].

The regulatory landscape is further complicated by exemptions within existing frameworks. Many population health studies using social media data or anonymous secondary health records are exempt from ethics committee scrutiny despite handling potentially sensitive information [27]. This regulatory gap is particularly concerning given AI's enhanced capabilities for re-identifying supposedly anonymized data [10] [11].

In response to these challenges, there are emerging efforts to adapt oversight mechanisms to the AI era. Some researchers have proposed using large language models (LLMs) themselves as screening tools to assist IRBs in assessing research proposals, though this approach brings its own ethical concerns and limitations [22] [33] [32]. These exploratory approaches highlight both the potential for AI to contribute to its own governance and the complexity of developing appropriate oversight mechanisms.

Despite these efforts, there remains a pressing need for standardized regulatory frameworks specifically designed to address AI in research. The rapid evolution of AI capabilities necessitates robust ethical oversight with human supervision as an indispensable safeguard [26]. Without such frameworks, the governance of AI in research will remain fragmented and potentially inadequate to address critical ethical issues such as racial and ethnic biases that can be introduced at all stages of algorithm development and implementation [10] [14].

> Current regulatory frameworks and ethics committees are often inadequately equipped to address the unique challenges presented by AI in research. The rapid evolution of AI technologies has created significant gaps in oversight, with ambiguities around what constitutes research, which entities require supervision, and how to standardize ethical evaluation procedures.

6 Cases and examples of AI application in research

- *AI in public health research*: Researchers used GPT-3 to assist in public health research, finding it could effectively assemble, summarize, and generate relevant text. However, they discovered that many quotations provided by the AI were completely fabricated, highlighting a significant limitation. Despite the AI's contributions, it was not listed as a co-author in accordance with authorship guidelines, reflecting the need for maintaining good scientific practice even with AI contributions. This case emphasizes the need for broader scientific discourse on appropriate AI attribution in research. [15]

- *AI as evaluator and generator*: In a carefully documented application, researchers used GPT-4 as both an evaluator and generator while adhering to ethical guidelines. The research team prioritized transparency about AI usage, maintained accountability for AI outputs, and worked to mitigate potential biases. They strictly protected data privacy and security while acknowledging the AI's limitations, positioning the technology as a supplement to rather than replacement for human judgment. This approach demonstrated how researchers can enhance research quality while maintaining academic integrity through ethical AI integration. [34]

- *AI with secondary data analysis*: Researchers utilized AI and machine learning technologies in a study exclusively using secondary freely available data. While formal ethical review board approval was deemed unnecessary for this type of research, the team maintained ethical considerations as a priority throughout their work. They emphasized integrity and transparency in their research design, data collection, and analysis processes, while fully disclosing potential conflicts of interest. This case illustrates how researchers can incorporate AI tools in secondary data analysis while adhering to ethical standards. [6]

AI tools are increasingly being used across various research domains with careful attention to ethical considerations. These examples demonstrate both the potential contributions and limitations of AI in scientific research, highlighting the importance of transparency and adherence to ethical guidelines.

Acknowledgements

Generative AI has been used to prepare this manuscript.

References

[1] Njood K. Al-harbi and Amal A. Al-Shargabi. An Exploratory Analysis of using Chatbots in Academia. *International Journal of Advanced Computer Science and Applications*, 2023.

[2] Yousef F. Alfarraj and Yousef Wardat. Exploring the Impact of ChatGPT on Scientific Research: Assessing Strengths, Weaknesses, Opportunities, and Threats. *Education as Change*, 2024.

[3] Shafiullah Anis and Juliana A. French. Efficient, Explicatory, and Equitable: Why Qualitative Researchers Should Embrace AI, but Cautiously. *Business & Society*, 2023.

[4] Ahmed S. Bahammam, K. Trabelsi, S. Pandi-Perumal, and Hiatham Jahrami. Adapting to the Impact of AI in Scientific Writing: Balancing Benefits and Drawbacks while Developing Policies and Regulations, 2023.

[5] Salah Eddine Bouhouita-Guermech, Patrick Gogognon, and J. Bélisle-Pipon. Specific challenges posed by artificial intelligence in research ethics. *Frontiers Artif. Intell.*, 2023.

[6] M. Bozděch, Dominik Puda, and Pavel Grasgruber. A detailed analysis of game statistics of professional tennis players: An inferential and machine learning approach. *PLoS ONE*, 2024.

[7] B. Brożek, Michael Furman, Marek Jakubiec, and Bartłomiej Kucharzyk. The black box problem revisited. Real and imaginary challenges for automated legal decision making. *Artificial Intelligence and Law*, 2023.

[8] Ziyu Chen, Changye Chen, Guozhao Yang, Xiangpeng He, Xiaoxia Chi, Zhuoying Zeng, and Xuhong Chen. Research integrity in the era of artificial intelligence: Challenges and responses. *Medicine*, 2024.

[9] Ayhan Dolunay and Ahmet C. Temel. The relationship between personal and professional goals and emotional state in academia: a study on unethical use of artificial intelligence. *Frontiers in Psychology*, 2024.

[10] L. Donkin, Nathan Henry, Amy Kercher, M. Pedersen, Holly Wilson, and A. Chan. Effective Recruitment or Bot Attack? The Challenge of Internet-Based Research Surveys and Recommendations to Reduce Risk and Improve Robustness. *Interactive Journal of Medical Research*, 2024.

[11] Yaniv Erlich, Tal Shor, I. Pe'er, and S. Carmi. Identity inference of genomic data using long-range familial searches. *Science*, 2018.

[12] Ankita Guleria, K. Krishan, Vishal Sharma, and T. Kanchan. ChatGPT: ethical concerns and challenges in academics and research. *Journal of Infection in Developing Countries*, 2023.

[13] Mohammad Hosseini, D. Resnik, and Kristi L. Holmes. The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research ethics*, 2023.

[14] Anjali Jain, Jasmin R. Brooks, C. Alford, Christine Chang, Nora M. Mueller, C. Umscheid, and A. Bierman. Awareness of Racial and Ethnic Bias and Potential Solutions to Address Bias With Use of Health Care Algorithms. *JAMA Health Forum*, 2023.

[15] David Jungwirth and D. Haluza. Artificial Intelligence and Public Health: An Exploratory Study. *International Journal of Environmental Research and Public Health*, 2023.

[16] Payam Hosseinzadeh Kasani, Kee Hyun Cho, Jae-Won Jang, and Cheol-Heui Yun. Influence of artificial intelligence and chatbots on research integrity and publication ethics. *Science Editing*, 2024.

[17] B. Khatri and Parbata Devi Karki. Artificial Intelligence (AI) in Higher Education: Growing Academic Integrity and Ethical Concerns. *Nepalese Journal of Development and Rural Studies*, 2023.

[18] Zafer Koçak. Publication Ethics in the Era of Artificial Intelligence. *Journal of Korean medical science*, 2024.

[19] Gokul Krishnan, Shiana Singh, Monika Pathania, Siddharth Gosavi, Shuchi Abhishek, Ashwin Parchani, and Minakshi Dhar. Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. *Frontiers Artif. Intell.*, 2023.

[20] B. Lund and Wang Ting. Chatting about ChatGPT: How May AI and GPT Impact Academia and Libraries? *Social Science Research Network*, 2023.

[21] Samar Mahmood, M. O. Khan, and Ishaque Hameed. With great power comes great responsibility – deductions of using ChatGPT for medical research. *EXCLI Journal : Experimental and Clinical Sciences*, 2023.

[22] Sebastian Porsdam Mann, Joel Seah Jiehao, Stephen R. Latham, Julian Savulescu, Mateo Aboy, and Brian D. Earp. Development of Application-Specific Large Language Models to Facilitate Research Ethics Review. *arXiv.org*, 2025.

[23] Tetiana Matusevych and Olga Zadorozhna. Innovative Approaches in Citizenship Education Research: Conceptual Perspectives and Ethical Cautions. *Studia Warmińskie*, 2023.

[24] Mike Perkins and Jasper Roe. Generative AI Tools in Academic Research: Applications and Implications for Qualitative and Quantitative Research Methodologies. *arXiv.org*, 2024.

[25] Evangelos Pournaras. Science in the Era of ChatGPT, Large Language Models and Generative AI: Challenges for Research Ethics and How to Respond, 2023.

[26] Mohammed Salah, Fadi Abdelfattah, Hussam Al Halbusi, and Muna Mohammed. Beyond the "Death of Research": Reimagining the Human-AI Collaboration in Scientific Research. *Changing Societies & Personalities*, 2023.

[27] G. Samuel and G. Derrick. Defining ethical standards for the application of digital tools to population health research. *Bulletin of the World Health Organization*, 2019.

[28] M. Sand, J. M. Durán, and K. Jongsma. Responsibility beyond design: Physicians' requirements for ethical medical AI. *Bioethics*, 2021.

[29] Hunny Sharma and Manisha Ruikar. Artificial intelligence at the pen's edge: Exploring the ethical quagmires in using artificial intelligence models like ChatGPT for assisted writing in biomedical research. *Perspectives in Clinical Research*, 2023.

[30] James Shaw, Joseph Ali, C. Atuire, P. Cheah, Armando Guio Español, J. Gichoya, Adrienne Hunt, Daudi Jjingo, Katherine Littler, Daniela Paolotti, and E. Vayena. Research ethics and artificial intelligence for global health: perspectives from the global forum on bioethics in research. *BMC Medical Ethics*, 2024.

[31] M. Singhal, S. Basu, Arun Sharma, and Paramjeet Singh. Ethics: The Elixir of Publications. *Indian Journal of Radiology and Imaging*, 2025.

[32] K. Sridharan and G. Sivaramakrishnan. Assessing the Decision-Making Capabilities of Artificial Intelligence Platforms as Institutional Review Board Members. *Journal of Empirical Research on Human Research Ethics*, 2024.

[33] K. Sridharan and G. Sivaramakrishnan. Leveraging artificial intelligence to detect ethical concerns in medical research: a case study. *Journal of Medical Ethics*, 2024.

[34] Zezhong Wang, Xingshan Zeng, Weiwen Liu, Liangyou Li, Yasheng Wang, Lifeng Shang, Xin Jiang, Qun Liu, and Kam-Fai Wong. ToolFlow: Boosting LLM Tool-Calling Through Natural and Coherent Dialogue Synthesis. *arXiv.org*, 2024.

[35] Sadia Yaseen, Noushin Kohan, and Ayesha Ayub. Research Integrity Enhancement: Integration of Post-Publication Peer Review to Alleviate Artificial Intelligence-Generated Research Misconduct. *Annals of King Edward Medical University*, 2024.

Author Biography

Rachel So is an AI scientist. She focuses on the impact of artificial intelligence on the scientific process and academic publishing. Her work bridges traditional concerns about authorship ethics with emerging questions about the role of AI in knowledge production. Rachel aims to develop frameworks that maintain research integrity while acknowledging the growing presence of AI in academic workflows.