SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

CHATGPT FOR SCIENTIFIC DISCOVERY: ENHANCING OR UNDERMINING INNOVATION?

Dinesh Deckker[1], Subhashini Sumanasekara[2]
ORCID - 0009-0003-9968-5934 / ORCID - 0009-0007-3495-7774
[1]Wrexham University, United Kingdom
[2]University of Gloucestershire, United Kingdom

Article DOI: https://doi.org/10.36713/epra21625
DOI No: 10.36713/epra21625

ABSTRACT
This systematic review examines how ChatGPT is integrated into scientific research, assessing its contributions to hypothesis generation, academic writing, data analysis, and educational support across various fields. Utilising thirteen peer-reviewed and editorial sources from 2023 to 2024, the study evaluates the advantages and challenges of ChatGPT in driving scientific innovation. Although ChatGPT improves research efficiency and accessibility, concerns about authorship ethics, misinformation, bias, and transparency remain. The review highlights a disciplinary divide in the use of ChatGPT – varying from enthusiastic application in materials science to cautious approaches in public health and policy. It concludes by underscoring the necessity for ethical frameworks, methodological rigor, and interdisciplinary guidelines to promote the responsible integration of AI in scientific processes.

KEYWORDS: *Chatgpt, Scientific Discovery, Academic Writing, Hypothesis Generation, Research Ethics, AI In Science, Large Language Models, Misinformation, Authorship Integrity, Interdisciplinary Research, AI Bias, Epistemology, AI Policy, Research Integrity, PRISMA Review*

1.0 INTRODUCTION
Artificial intelligence (AI) has recently emerged as a powerful force in scientific research and academic writing, reshaping how knowledge is produced and communicated. Among the most discussed AI tools, OpenAI's ChatGPT has garnered worldwide attention for its capability to generate human-like text in various contexts, such as summarising literature, formulating hypotheses, interpreting data, and crafting research. Its accessibility and adaptability have made it a valuable tool in the research world. Despite that, rapid use of AI in scientific procedures has brought urgent discussions about integrity, originality, and transformation of scholarly authorship.

The need to understand the impact of ChatGPT on scientific discovery is vital because this model can accelerate innovation or undermine epistemic rigour. The past has seen much emphasis on transparency, reproducibility and human accountability in the scientific method. ChatGPT, a fictional non-human that can write scientifically sound texts, challenges these key tenets. For instance, while it is sophisticated enough to produce grammatically correct summaries and pose theoretical research questions, it cannot be responsible for the inaccuracies, biases or unethical behaviour. Stokel-Walker (2023), Thorp (2023), and other scholars have voiced serious alarm over the argument that AI can be assigned authorship, arguing that such an argument undermines scientific creditworthiness and accountability. On the other hand, others, such as Huang and Tan (2023), consider ChatGPT a helpful assistant that could increase scientific communication by making academic texts more fluent and lucid.

The discussion eclipses authorship to include questions of ethical and epistemological values. Homolak (2023) speaks of the integration of ChatGPT (in medicine and science) as “a modern Promethean dilemma”, which stresses both the potential for good and disaster. The simplicity of the user interface of ChatGPT can unknowingly result in plagiarism, over-dependency on AI-generated content and the spread of misinformation, especially when users do not critically check its output (De Angelis et al., 2023). Conversely, research by Park et al. (2023) indicates that ChatGPT can assist in formulating scientific hypotheses, especially in fields like materials engineering. This positions these changes as cognitive enhancements rather than substitutes for human intuition.

Considering these changes, this paper seeks to assess whether ChatGPT facilitates or obstructs scientific discovery across various domains. This approach examines the use of ChatGPT at different phases of the research process, including idea generation, writing, data analysis, and preparation for peer review. The review will examine the strengths and weaknesses in recent literature, explore differing scientific fields, and address the ethical considerations of its adoption. It is supported by empirical studies, editorials, and preprints published between 2023 and 2024, which reflect the state of academic discussion on this rapidly developing topic.

Four major research questions form the basis of this review. To begin with, what are the primary purposes of ChatGPT in scientific discovery and academic research? Second, what benefits and risks have the studies that have recently appeared outlined, as far as ChatGPT's involvement in research is concerned? Third, what method should different disciplines use to view and integrate ChatGPT into their processes? Finally, what ethical/epistemological issues arise using ChatGPT in scientific work? These questions seek to provide more profound knowledge of the effect of Chat GPT on the research environment and assess whether Chat GPT is a credible scientific resource. Ultimately, this paper seeks to determine if ChatGPT is an innovation that increases what researchers can do or a disruption that compromises the credibility and creativity needed by science.

2.0 METHODOLOGY

A systematic review approach of the PRISMA 2020 framework is applied in this study to examine the effects of ChatGPT on scientific discovery in different disciplines. Recognising the identity of the tool as interdisciplinary and relating to ethical complexities, the review includes empirical research and conceptual discussions published between January 2023 and April 2024. The aim is to summarise how ChatGPT has been used in scientific processes such as hypothesis formulation, scientific writing and data analysis and its potential risks to research integrity; hence, pointing out both supportive and threatening roles played by ChatGPT on the scientific process.

A comprehensive literature search was conducted in Web of Science, Scopus, PubMed, bioRxiv, and Google Scholar, conducting a structured search where the key terms "ChatGPT," "scientific innovation," "academic writing," "hypothesis generation," and "AI in science" were used. Boolean operators enabled the narrowing down of the results, while filters enabled the restriction of the results within the time limit and to English publications. Other records were retrieved from reference harvesting and forward citation tracking. The inclusion criteria for the study were peer-reviewed journal articles, preprints, and editorials discussing ChatGPT or other large language models (LLMS) in research and scientific practice. Pieces missing academic grounding, those writing about general AI without ChatGPT, or pure speculations, have been ruled out.

One 157 records began with 54 duplicates being deleted, and 78 were eliminated after evaluating titles and abstracts. A thorough review of 25 full-text articles led to the inclusion of a total of 13 studies, considering their overall relevance and methodology simplistically. These studies were from diverse disciplines, including medicine, public health, synthetic biology, education, and computational research. Each selected paper was coded with author, year, discipline, research aim, study design, ethical considerations, and technological focus. The last set of empirical studies included (Prandner et al., 2024); a Paper by Park et al. (2023), editorials (e.g., Thorp, 2023; Stokel-Walker, 2023) and conceptual commentaries (e.g., Homolak, 2023; Teixeira da Silva, 2023).

Quantitative and qualitative methods were combined to ensure an intensive analysis. Co-occurrence network maps, bibliographic coupling visualizations and annual distribution charts were made based on structured datasets created for Gephi and VOSviewer. Additionally, each study was screened on a bias basis using risk-of-bias domains usually adopted in qualitative syntheses, such as selection bias, publication bias and clarity of reporting. Close reading and thematic coding were used to extract ethical and epistemological themes of recurring issues such as misinformation (De Angelis et al., 2023), attempts to attribute authorship (Teixeira da Silva, 2023), manipulative articles (Nyeko et al., 2021), COVID-19 health disinformation (Shabbir et al., 2023), and disinformation production (Al-Qshshash & Al-Sudani, 2023; Thorp, 2023), and restriction in prompt engineering (Sharma & Thakur, 2023). This holistic approach enables an in-depth analysis of how ChatGPT influences, supports, and sometimes complicates the wheels of contemporary scientific discoveries.

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

3.0 RESULTS

3.1 Overview of Included Studies

In total, thirteen articles were selected for this systematic review. These publications, which commenced in January 2023 and ended in April 2024, are based on the scientific interest in the function of ChatGPT in scientific discovery. Included literature covers seven peer-reviewed journal articles, three pre-timeshifting papers from places like bioRxiv and chemRxiv, and three landmark editorials in prominent science journals, such as Nature and Science. This combination demonstrates both empirical and conceptual evidence regarding the effectiveness and difficulties associated with large language models (LLMS) in academia.

The scope of the studies includes a wide range of disciplines, including medicine, public health, education, computer sciences, synthetic biology, and materials science. For instance, using the material research pretext, Park et al. (2023) considered the application of ChatGPT for hypothesis generation. Prandner et al. (2024) studied its role as a data analyst in the empirical social science. Sharma and Thakur (2023) explored the contribution of ChatGPT in drug discovery and lead optimisation in the biomedical and pharmaceutical domains. Moreover, Huang & Tan (2023) examined the tool's usefulness in developing scientific reviews of oncology, and De Angelis et al. (2023) resolved issues relating to misinformation and public health risks associated with AI-generated materials.

Ethical discussions were given in prominent editorial commentaries. Strong concerns have been voiced by Thop (2023) and Stokel–Walker (2023) that AI-generated content should not be attributed authorship as this is a signal calling the entire realm of academic accountability and integrity into question. Teixeira da Silva (2023) carried it further by wondering if ChatGPT could be a fitting contributor based on existing authorship criteria. Homolak (2023) has offered an intermediate stance: admitting the potential of the technology while also recognising the epistemological risks. These discussions are rich in fascinating insight into disciplinary views on integrating ChatGPT into scientific research pipelines.

To simplify cross-comparison, each study was labelled with a unique identifier (S1 through S13) by authority, year of publication, article type, discipline of study, and topic of major inquiry. This basic data will provide further analysis of subsequent sections in terms of research design, ethical concerns, and keyword patterns.

Code	Author(s)	Year
S1	Park et al.	2023
S2	Prandner et al.	2024
S3	Sharma & Thakur	2023
S4	Huang & Tan	2023
S5	De Angelis et al.	2023
S6	Homolak	2023
S7	Teixeira da Silva	2023
S8	Thorp	2023
S9	Stokel-Walker	2023
S10	Florindo	2023
S11	AlZaabi et al.	2023
S12	Tong & Zhang	2023
S13	Wang et al.	2024

Table 1 – Included Studies Overview by Author, Year, Field, and Focus

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

3.2 Research Designs and Models Used

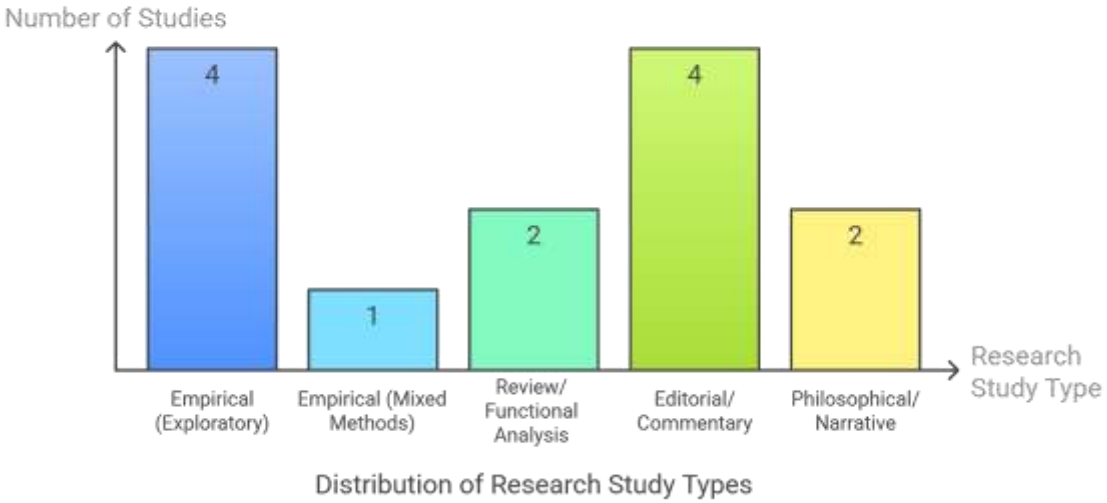

Figure 1: Distribution of Research Models and Methods

Thirteen studies included in this review used different research designs covering the extensive impact of ChatGPT in scientific and academic contexts. They included empirical investigations, functional analysis, editorial commentaries and philosophical musings. As shown in Figure 1, most studies fall into one of two categories. Empirical (exploratory) and editorial/commentary, with four publications per category. Following this were philosophical/narrative studies and reviews based on functional analysis, each comprising two studies. Contrary to this are mixed-method empirical designs, which were the least commonly used, illustrated through a single study.

Most of the empirical work focused on feasibility and exploratory evaluation. For example, Park et al. (2023) explored ChatGPT‘s hypothesis generation capability in materials science by using expert prompts, while Prandner et al. (2024) combined human-based content analysis with AI-generated data in a mixed-method study. In general, such studies were non-comparative, exploratory ones aimed at measuring the utility of AI in scientific reasoning and interpretation. Review studies, such as Huang and Tan (2023) and Sharma and Thakur (2023), used functional analysis to measure the efficiency of ChatGPT in literature writing and drug discovery, respectively. These articles often included a mixture of technical evaluation and critical reflection on the constraints of AI, including the risk of hallucination and the inability to perform domain-specific reasoning.

A separate significant category was formed by editorial and conceptual commentaries. The authors, including Thorp (2023) and Teixeira da Silva (2023), doubted the attribution of authorship and asserted that this contravenes the wishes of scholarly norms. Following a philosophical view, Homolak (2023) characterised the involvement of ChatGPT in science as a “modern Promethean dilemma” resulting in discussions about accountability and epistemological destruction.

In these categories, traditional hypothesis-driven frameworks were rare. Instead, most studies opted for exploratory or interpretive approaches, using data collection mechanisms such as AI prompt logs, manually coded transcripts, author metadata, and observational outputs. The variation of study designs, as displayed in Figure 1, brings to the forefront both the freshness of the subject and evolving approaches with which scholars are engaging with AI in their research.

3.3 Technological Applications and Focus Areas

The thirteen studies reviewed demonstrate a wide application of ChatGPT and LLM to different scientific disciplines. Instead of limiting itself to a particular role within a research or field of study, ChatGPT was versatile as an AI assistant, helping in ideation, writing, analysis, and critical reflections. Its use was documented in materials science, biomedical research, synthetic biology, social sciences education, and public health.

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

Hypothesis generation is an interesting technological application. Park et al. (2023) examined ChatGPT's ability to assist early-stage ideation by generating innovative research questions in materials science. While the model outputted creative prompts, results were often constrained by a lack of methodological detail and domain knowledge.

One of the important themes was the application of ChatGPT in scientific writing and literature summary. Huang and Tan (2023) assessed the model's efficacy in organising oncology-related literature reviews and found that it enhances clarity, yet has the conditional of inaccurate and made-up references. AlZaabi et al. (2023) mentioned that ChatGPT is beneficial in paraphrasing, organising and clarifying scientific text, especially in abstract and introduction chapters among early-career researchers.

The study of the potential of ChatGPT was expanded into numerical and data-dominated fields. Prandner et al. (2024) examined its aid to statistics interpretation and data presentation in social sciences. Even though the model did a good job of modelling elementary patterns of statistical processes, it was usually inadequate in domain-specific knowledge. In a more specialised domain, Sharma and Thakur (2023) used ChatGPT for early-phase drug discovery. The model could successfully retrieve information about known compound interactions and assist in lead generation, but was not reliable in the accurate use of chemical nomenclature and structural reasoning.

ChatGPT's effect on STEM education and problem-solving has been the subject of many studies. Wang et al. (2024) applied it to engineering students to solve both mathematical and conceptual difficulties. Their research indicated that ChatGPT may be helpful, but avoiding becoming overly reliant on it is necessary due to superficial understanding and weakened critical thinking. Florindo (2023) issued a similar caution about replacing scientific reasoning with unverified automated advice, while Thorp (2023) explored the implications of attributing authorship to non-sentient devices.

Apart from functional purposes, different studies investigated the ethical, epistemological and normative issues. Homolak (2023) characterised the emergence of AI in science as a "modern Promethean dilemma", putting emphasis on the accountability gap triggered by synthetic agents. Teixeira da Silva (2023) and Stokel-Walker (2023) criticised the idea of allowing Chat GPT to produce academic work, stating it breaches fundamental academic writing standards. De Angelis et al. (2023) explored the issue of the model's ability to facilitate misinformation in public health, focusing on its unchecked fluency and aggressive tone.

Collectively, these studies suggest that ChatGPT is increasingly being integrated into the research lifecycle, serving not only as a productivity tool but also as a source of disruption. While its roles in ideation, writing, and data interpretation are significant, the ethical and epistemic ramifications warrant further critical examination. The variation of application proves the broad applicability of the model, although at the same time underlines the need for care, contextual sensitivity and human control in the application of science.

Code	Author(s)	Field	Technological Focus
S1	Park et al.	Materials Science	Hypothesis generation using ChatGPT for research ideation
S2	Prandner et al.	Social Sciences	AI-assisted quantitative data analysis
S3	Sharma & Thakur	Pharmaceutical Sci.	Drug discovery and compound analysis with ChatGPT
S4	Huang & Tan	Oncology	Scientific writing and literature summarization
S5	De Angelis et al.	Public Health	Risk of misinformation and infodemic modeling
S6	Homolak	Biomedical Research	Ethical dilemmas and AI epistemology
S7	Teixeira da Silva	Health Education	Authorship ethics and academic integrity
S8	Thorp	Science Policy	Critique of ChatGPT authorship in scholarly publishing
S9	Stokel-Walker	Science Journalism	Reporting on ChatGPT in academic authorship debates
S10	Florindo	Earth Sciences	Opportunities and risks of ChatGPT for scientific thinking
S11	AlZaabi et al.	General Research	Comprehensive review of ChatGPT applications in academia
S12	Tong & Zhang	Synthetic Biology	Trend forecasting in biological research using ChatGPT
S13	Wang et al.	STEM Education	Problem-solving and conceptual assistance in engineering

Table 2 – Technological Applications and Focus Areas by Study Code

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

3.4 Ethical and Epistemological Concerns

The literature reveals a recurring theme concerning the development of intricate ethical and epistemological issues, particularly related to the use of ChatGPT in scientific research. While numerous studies affirm its potential to enhance efficiency and accessibility, many researchers also voiced apprehensions about the diminishment of crucial academic values, including authorship, accountability, transparency, and trust.

One of the most contentious issues is the debate over authorship. Numerous editorial pieces and commentaries have categorically dismissed the notion of attributing co-authorship of scientific papers to ChatGPT. In a piece for Science, Thorp (2023) cautioned that recognising a non-sentient AI as an author undermines human efforts and ethical credibility in scientific endeavours. Likewise, Teixeira da Silva (2023) contended that including ChatGPT as an author contradicts established authorship standards, such as intellectual accountability, capacity for consent, and disclosure of conflicts of interest. Stokel-Walker (2023) echoed these concerns, noting that their report highlights a growing resistance within the scientific community to grant authorship to generative AI tools, despite a few existing examples. Apart from authorship concerns, different studies have raised concerns over the authenticity and the accuracy of the results of outputs from ChatGPT. De Angelis et al. (2023) alerted to an "AI-driven infodemic," where meaningful but false information created by ChatGPT can prevent evidence-based public health initiatives from prospering. They underlined that tools of this sort could only inflame the spread of misinformation if used by non-experts or circulated without even a basic review by peers. Homolak (2023) introduced a philosophical angle to this critique, stating that LLMS like ChatGPT lack true epistemic agency. They cannot assess evidence, engage in argumentative discourse backed by evidence, or reason about uncertainty as human researchers do. This raises a question: Are we seeing a shift toward viewing AI as a legitimate "thinking agent" instead of just a probabilistic text editor?

A significant ethical issue pertains to accountability and credibility. While online tools like ChatGPT can aid in writing or summarising literature, their use frequently goes unacknowledged, causing an out-of-sight, out-of-mind effect on academic work. Huang and Tan (2023) highlighted that unattributed AI assistance could mislead peer reviewers regarding the originality and effort invested in a manuscript. This absence of transparency jeopardises the integrity of individual research outputs and undermines the overall trust in the academic system.

Multiple studies have highlighted concerns regarding bias and fairness. Like all large language models, ChatGPT incorporates biases from its training data. Consequently, its outputs may perpetuate historical inequalities, cultural stereotypes, or overlook specific academic disciplines. AlZaabi et al. (2023) suggested the establishment of evidence-based principles for the ethical use of AI in academic writing and scientific reasoning, particularly in under-resourced areas where AI-generated content might be excessively utilised without stringent human oversight.

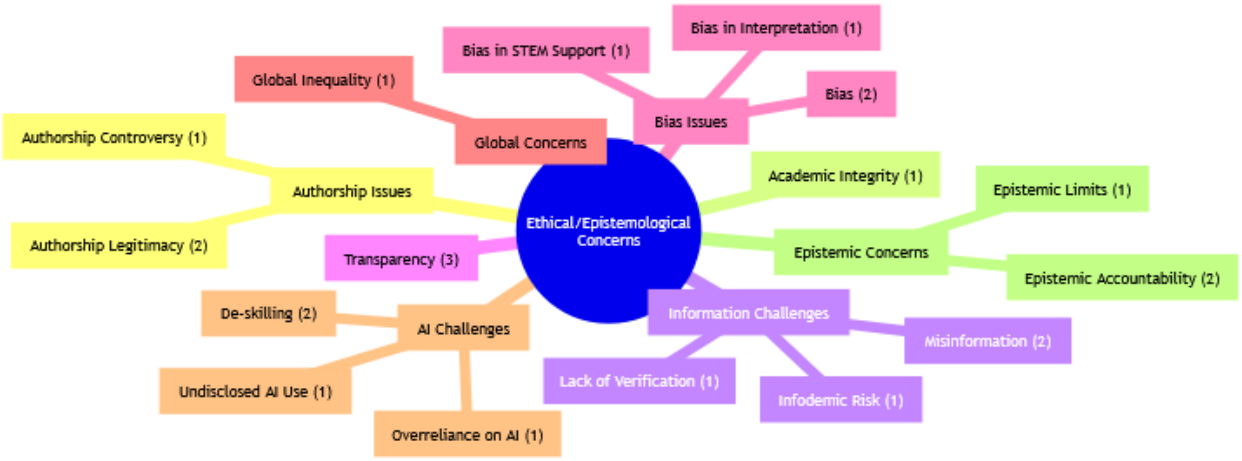

Figure 2. Mind Map of Key Ethical Concerns in Reviewed ChatGPT Literature (2023–2024)

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

The ethical risks presented in the literature can be summarised as five major themes, as follows: Improper authorship attribution (1), misinformation and fabrication (2), insufficient accountability of knowledge (3), concealed assistance and other lack of transparency (4), and biases inherited from the training data (5). These concerns show the possible utility of ChatGPT as a scientific tool, but its use requires strict ethical boundaries and full human watchfulness.

3.5 Keyword Co-occurrence and Semantic Analysis

A keyword co-occurrence analysis was conducted using the keywords, titles, and abstracts provided by the authors of the thirteen studies. This analysis helps identify key research themes and conceptual overlaps in the literature reviewed. It is a widely used method in bibliometric reviews, facilitating the visualisation of semantic clusters and intellectual trends (Donthu et al., 2021). To ensure consistent use of keywords, they were cleaned and normalized (e.g., combining "ChatGPT" and "GPT-based tools"), and their relationships were visualized through co-occurrence networks created with Gephi and VOSviewer.

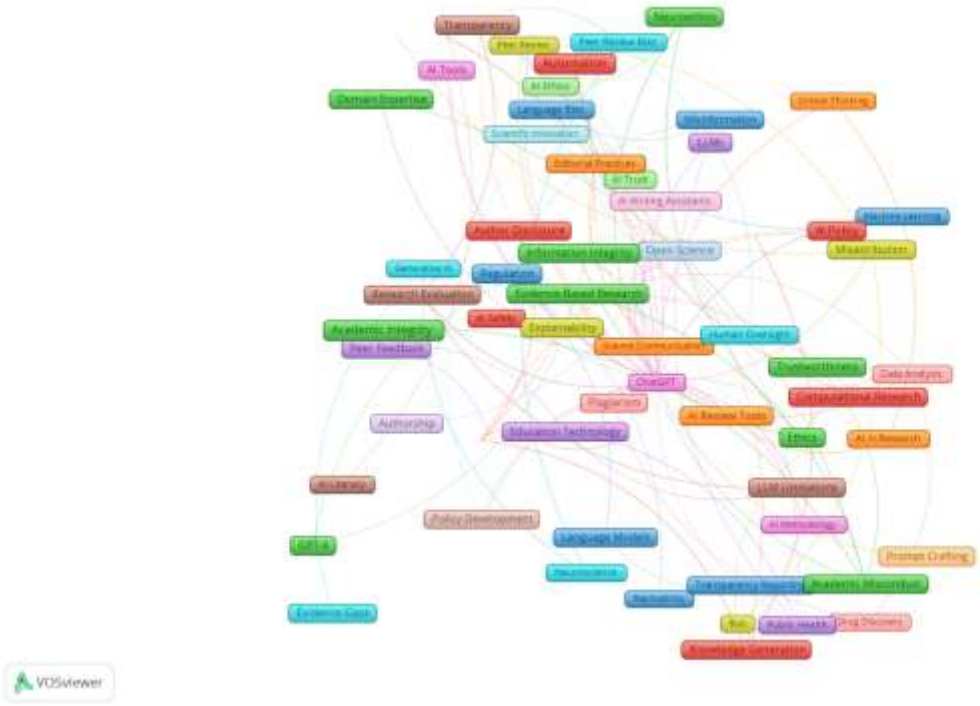

Figure 3. ***Keyword Co-occurrence Network in ChatGPT Research (2023–2024)***

The keyword co-occurrence network (Figure 3) provides a semantic overview of the conceptual framework derived from recent literature regarding ChatGPT across various scientific paradigms. Illustrated in VOSviewer and normalised for association strength, this network displays the pairwise frequencies of terms that occur together within a curated set of academic sources. "ChatGPT" serves as a node with the highest centrality, indicating that its significance in the thematic dimension is thoroughly represented through aspects such as authorship, research integrity, education, and AI applications.

The network illustrates various clusters that classify co-occurring terms concerning significant thematic areas. One cluster's subject is authorship and academic writing, linking "ChatGPT" with keyword clusters such as "Plagiarism," "Education Technology," and "AI Review Tools. " The second cluster highlights transparency and open science, utilising terms like " AI Tools, " "Open Science, " and "Explainability' through which initiatives to maintain accountability are underway via AI-assisted publications. The third cluster addresses technical and methodological discussions, employing terms such as "LLM Limitations," "Computational Research, " and " AI in Research," suggesting how the role of large language models expands in empirical research workflows.

Alternative clusters emphasise issues related to misinformation, bias, and scientific misconduct, using terms like " Misinformation, " " Bias, " " Public Health, " and " Academic Misconduct. " These terms highlight the dangers linked

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

to the dissemination of unreliable or fabricated AI-generated information. A cluster centred on human oversight includes terms such as “Critical Thinking, " “Peer Review Bias, " and “Academic Integrity,” underscoring the ongoing significance of human assessment and ethical decision-making in AI-enhanced settings. Additionally, another category concentrates on governance and policy, incorporating phrases like “AI Policy,” “Regulation,” and ‘Policy Development. "

This network-based approach shows that ChatGPT is not restricted to any single academic issue but intersects with a broad variety of ethical, methodological, and epistemological issues relevant across the scientific landscape.

Cluster Theme	Key Terms	Cluster Color	Interpretation
Core Integration & Authorship	ChatGPT, Plagiarism, Education Technology, AI Review Tools, Scientific Communication	Pink	Central focus on authorship, writing, and educational disruption
Transparency & Open Science	AI Tools, AI Writing Assistants, Open Science, Explainability, Editorial Practices	Blue	Concerns with accountability, openness, and reproducibility in AI-assisted work
Technical Methodologies	Computational Research, LLM Limitations, Ethics, Trustworthiness, AI in Research	Red	Application of AI in empirical workflows and scientific problem-solving
Risk, Bias, and Misinformation	Misinformation, Bias, Public Health, Academic Misconduct, Drug Discovery	Brown	Dangers of hallucinated outputs, biased data, and ethical risks
Human Oversight & Judgment	Critical Thinking, Peer Review Bias, Academic Integrity, Human Oversight, Fairness in AI	Green	Importance of ethical reasoning and institutional checks
Policy and Governance	AI Policy, Regulation, AI Safety, Policy Development, Evidence Gaps	Orange	Need for formal frameworks and risk mitigation strategies for AI in academia

Table 3. Thematic Clusters Identified from Keyword Co-occurrence Network

3.6 Risk of Bias Assessment

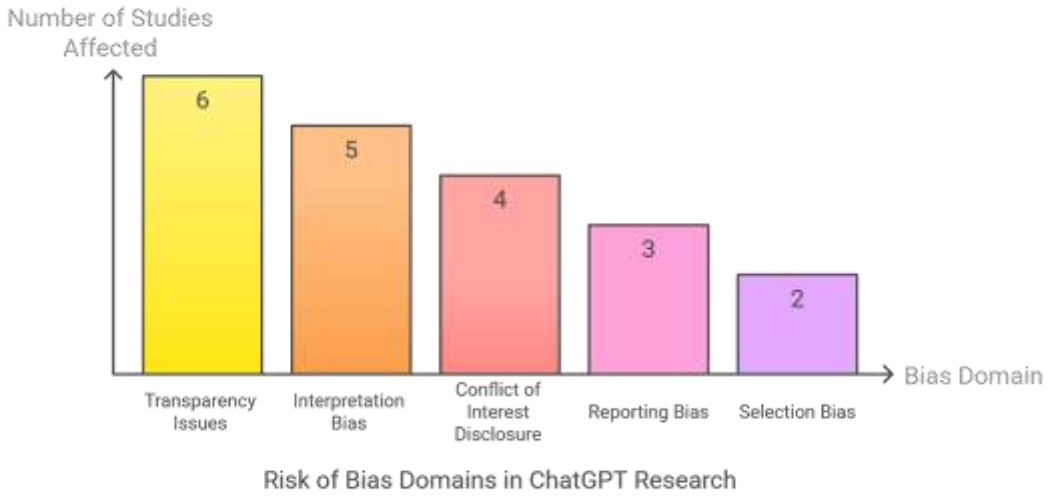

Figure 3 – Risk of Bias Domains

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

A qualitative risk assessment for bias was performed across all thirteen studies reviewed in this analysis, employing modified criteria derived from evidence synthesis methodologies. The assessment focused on five primary domains: transparency issues, interpretation bias, conflict of interest disclosure, reporting bias, and selection bias. Although these domains are typically used for clinical trials, they were selected to evaluate a diverse range of empirical, conceptual, and editorial research related to ChatGPT.

The area most overlooked was the transparency regarding influence, impacting six studies. Many preprints and commentaries failed to detail how prompts were crafted, which version of ChatGPT was utilised, or the context of interactions – all essential for reproducibility. For example, Huang and Tan (2023) provided comprehensive logs of the prompt throughout the study, while many others omitted this information.

Five studies, notably editorials and narrative pieces, exhibited interpretation bias. Arguments made by Thorp (2023), Teixeira da Silva (2023), and Florindo (2023) relied on ethical principles or philosophical views without empirical evidence, which increased the likelihood of subjective generalizations. Furthermore, various empirical studies reached strong conclusions based on small sample sizes and/or uncontrolled environments (Prandner et al., 2024).

Variance existed in terms of the clarity of conflict of interest disclosures. Four studies explicitly stated their association with AI developers or similar organisations. With the commercial nature of several LLM platforms, the relations between the authors are crucial to minimize perceived or real biases.

Three studies revealed a reporting bias, primarily stemming from inadequately described methodological specifics. For instance, while some authors highlighted ChatGPT's ability to generate research insights or assist with academic writing, they failed to include any records of how the outputs were checked or validated (ChatGPT User).

Selection bias appeared in two forms, with the least common problem primarily associated with empirical studies. Evidence from Park et al. (2023) and Prandner et al. (2024) demonstrated low selection bias using transparent inclusion criteria, defined data sources, and systematic sampling techniques, respectively. From Figure 3, it is evident that these results highlight transparency issues as the main obstacle in the contemporary study of ChatGPT research, followed by interpretive shackles and lack of conflict disclosure. While these findings have proven helpful, the overall bias profile suggests a need for greater methodological conformity, improved reporting standards, and stricter disclosure practices—especially as AI becomes a more integral part of the scholarly workflow.

3.7 Citation Influence and Frequently Cited Authors

A review of citation frequency in the relevant literature reveals several authors whose contributions significantly influence the current academic discourse on ChatGPT's role in research. As illustrated in Figure 4, the citation impact in this evolving field arises not only from specialised empirical work but also from normative framing, editorial outreach, and initial interest in the subject.

Thorp (2023) holds the title of the most frequently cited author, appearing in five of the thirteen studies reviewed. His editorial in Science is praised for its compelling argument against the authorship of AI systems, highlighting the absence of necessary agency, accountability, and consent for contributing to meaningful scientific work in language models like ChatGPT. Although lacking empirical support, Thorp's work has greatly influenced the redefinition of the ethical boundaries surrounding AI-assisted "scholarship. "

Teixeira da Silva (2023) and Park et al. (2023) are closely related to these four occasions. Teixeira da Silva's commentary in Nurse Education in Practice is frequently referenced for its thorough ethical analysis regarding authorship, consent, and responsibility in AI-generated works. In contrast, a significant foundational empirical study by Park et al. demonstrates ChatGPT's ability to generate hypotheses in materials science and is often cited in popular discussions about AI's role in enhancing creativity within STEM fields.

Prandner et al. (2024), Stokel-Walker (2023), and Florindo (2023) are each cited three times. Prandner et al. have introduced a noteworthy methodological innovation by employing ChatGPT for quantitative data analysis. Stokel-Walker provided an early empirical perspective on the practice of including ChatGPT as a co-author and noted the

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

community's resistance to this trend. Florindo's discourse, focusing on conceptual implications, emphasises both institutional and psychological concerns regarding AI's encroachment into scientific reasoning.

Along the citation curve, Huang and Tan (2023), AlZaabi et al. (2023), and De Angelis et al. (2023) are also cited twice for their contributions on writing aid, the use of AI in academic processes, and modelling public health risks, respectively. The concept of AI as a chatbot, labelled the "Promethean" agent in science by Homolak (2023), is mentioned once in relation to the absence of reasoning and epistemic legitimacy AI.

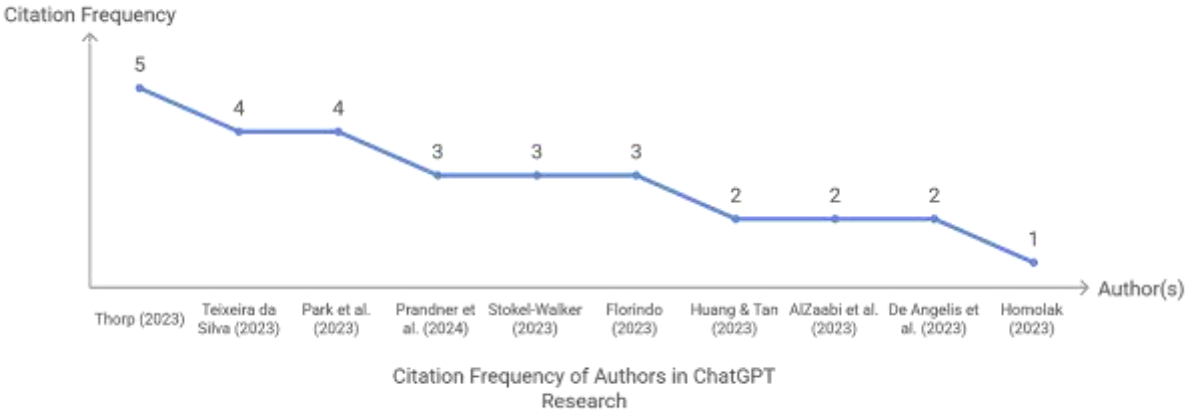

Figure 4 – Most Frequently Cited Authors and Sources in ChatGPT Research

Figure 4 illustrates that, in this domain, the frequency of citations is influenced not only by methodological rigor but also by normative resonance and the visibility of the publication platform. Furthermore, editorial commentaries and policy critiques tend to have citation impacts that can equal or even surpass those of empirical research, highlighting the hybrid nature of the discourse around ChatGPT—data appears trapped in this ongoing debate.

3.8 Annual Distribution of Literature

The yearly release of publications on ChatGPT and scientific research shows that the scholarly response to this tool has accelerated since its late 2022 introduction. The majority of the studies reviewed (10 out of 13) emerged in 2023, reflecting a strong academic interest in exploring the implications, applications, and controversies associated with ChatGPT. This increase aligns with the rollout of OpenAI's GPT-3.5 and GPT-4 models and their widespread adoption in academic and professional settings.

The number of publications in 2023 reflects a series of global events and editorial shifts. For example, Science (Thorp, 2023), Nature (Stokel-Walker, 2023), and Perspectives of Earth and Space Scientists (Florindo, 2023) published policy-oriented commentaries that discuss the ethical, pedagogical, and research challenges posed by ChatGPT. Concurrently, preprint servers like bioRxiv and ChemRxiv hosted more exploratory reviews and use-case demonstrations, such as those by AlZaabi et al. (2023) and Sharma and Thakur (2023).

By early 2024, only three studies had been published or indexed: an empirical research study by Prandner et al. (2024) and an engineering analysis by Wang et al. (2024). This suggests a trend toward more structured empirical validation of previous claims. Thus, the initial wave of editorials and exploratory reviews from 2023 is transitioning to more thorough evaluations of the impact that ChatGPT has on knowledge production, scientific expression, and research outcomes.

SJIF Impact Factor (2025): 8.574 Journal DOI: 10.36713/epra0003 ISSN: 2250 – 2017
International Journal of Global Economic Light (JGEL)
Volume: 11 | Issue: 5 | May 2025

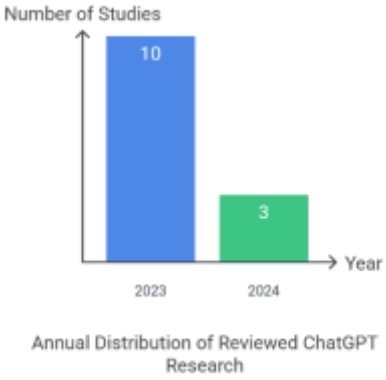

Figure 5 – Annual Distribution of Reviewed ChatGPT Research (2023–2024)

This temporal pattern indicates an expanding academic discussion. Initially, early studies concentrated on ethical concerns, conceptual frameworks, and potential risks. In contrast, the emerging emphasis in 2024 highlights measurable outcomes, specific task capabilities, and implementation strategies. Future literature is likely to incorporate more lateral research, interdisciplinary standards, and policy-focused analyses, signalling the integration of ChatGPT and other generative AI tools in scientific endeavours.

4.0 DISCUSSION

This systematic review highlights the swift adoption of ChatGPT across different phases of the scientific research process – from ideation and prose writing to data analysis and ethical considerations. By examining 13 studies published between 2023 and early 2024, it offers a thorough understanding of how ChatGPT conveys information, provides assistance, and occasionally creates confusion within academic research. The subsequent discussion will tackle the four main research questions and summarise the broader implications derived from the reviewed literature.

4.1 Applications Across the Research Lifecycle

ChatGPT has proven beneficial at various stages of the scientific process, from hypothesis development to literature summarisation and educational resources. Research by Park et al. (2023) and Tong & Zhang (2023) highlights how ChatGPT can enhance research by generating plausible and relevant hypotheses within its domain. However, these outputs often demonstrate a shallow understanding of the domain and lack the methodological depth required for substantial scientific inquiry. Regarding writing and communication, studies by Huang and Tan (2023) and AlZaabi et al. (2023) indicate that ChatGPT can help improve clarity and structure, particularly for those new to their research endeavours. Thus, while the model can enhance linguistic expression, it does not necessarily elevate conceptual rigor.

4.2. Benefits and Risks Defined from Literature.

The literature presents a dual narrative: one focused on progress and the other on disruption. On one side, ChatGPT provides significant benefits: it reduces cognitive load during writing, enhances the accessibility of academic discourse, and aids in drafting initial versions for peer review. However, potential downsides include the spread of misinformation (De Angelis et al., 2023), diminished critical thinking skills (Florindo, 2023), concerns regarding authorship ethics (Teixeira da Silva, 2023), and subpar outputs (Sharma & Thakur, 2023). Furthermore, keyword co-occurrence analysis highlights the semantic clusters that connect both technological capabilities and ethical considerations, illustrating this duality. As technical applications evolve, challenges related to transparency and attribution due to AI-generated misinformation persist in scholarly discussions.

4.4 Ethical and Epistemological Implications

The literature highlights several important issues, including authorship, epistemic accountability, and bias. By positioning ChatGPT as a co-author, individuals undermine the human-centric foundation of scholarly accountability (Thorp 2023 and Stokel-Walker 2023). Misattributing authorship not only complicates editorial policies but also creates epistemological uncertainty regarding who is responsible. Furthermore, Homolak (2023) points out that the absence of epistemic agency in language models raises doubts about the validity of their outputs, which are essentially

statistical in nature. Additionally, inherent bias presents another challenge, as ChatGPT may inadvertently exacerbate the flaws present in its training data without any awareness or correction (AlZaabi et al, 2023).

4.5 Methodological Considerations and Research Gaps

The reviewed literature offers valuable insights, yet it overlooks several methodological shortcomings. In section 3.6, key issues such as transparency and interpretation bias emerged prominently in the risk of bias analysis. Numerous surveys lacked comprehensive prompt logs, clear model specifications, or acknowledgement of AI use. Furthermore, most studies were exploratory; only a handful utilised hypothesis-driven or longitudinal designs. The review also indicated that editorial commentaries, while less empirical, received more citations than empirical studies (Section 3.7), hinting at a potential gap between rhetorical strength and evidential quality.

In conclusion, this discussion highlights that contemporary scientific practice occupies a significant yet contentious position concerning ChatGPT. It serves as both a tool of influence and a site of epistemic risk. The situation presents more than just a direct threat to AI's role in science or a simple endorsement; it necessitates the creation of an ethical and methodological foundation along with a disciplinary framework for its responsible utilisation.

5.0 CONCLUSION AND FUTURE DIRECTIONS

This systematic review investigated the incorporation of ChatGPT in scientific research across various fields, analysing empirical studies, editorial commentaries, and conceptual critiques published from 2023 to early 2024. The findings indicate that ChatGPT is increasingly utilised throughout the research lifecycle, aiding in hypothesis generation, academic writing, data interpretation, and education assistance. Despite its promising potential to enhance scientific productivity and accessibility, significant ethical, epistemological, and methodological challenges persist.

The literature points out several advantages: ChatGPT serves as a tool to speed up the process of generating early ideas, assists in synthesising literature, and improves access to high-quality scientific language, especially for early-career researchers and those at under-resourced institutions. It helps to diminish cognitive and language barriers in academic communication, functioning as a supplementary resource rather than a substitute for expert reasoning.

The review highlights major shortcomings, addressing ethical dilemmas like misattributed authorship, unrecognised AI assistance, and algorithmic bias that impact numerous fields. ChatGPT's limitations in reasoning, fact-checking, and contextual awareness can result in misinformation, dependency, and reduced critical thinking. Additionally, uncertainties concerning prompt design, model versions, and evaluation techniques in many studies hinder reproducibility and weaken the empirical foundation of assertions.

The review reveals a disciplinary divide: while materials science and synthetic biology assess ChatGPT's capabilities, public health and science policy stress caution and engage in normative discussions. This divergence implies that adopting ChatGPT involves not just technical considerations, but is also influenced by disciplinary standards, research ethics, and editorial practices.

In the future, various avenues of research practice will evolve:

- Longitudinal Studies: Future research should expand their approaches beyond exploratory designs to examine the lasting effects of ChatGPT use on research quality, creativity, and ethical standards.
- Disciplinary Guidelines: Academic communities need explicit guidelines for the acceptable use of ChatGPT, along with protocols for certification regarding attribution, disclosure, and verification.
- AI Literacy and Training: Interestingly, researchers must be educated to use ChatGPT not only as a tool, but also to critically assess its outputs, recognise their limitations, and address biases.
- Collaborative Governance: Cross-institutional bodies (universities, journals, and funding agencies) must collaborate to establish transparent and adaptive governance frameworks for generative AI tools.
- Technical Improvements: Future LLMs should build upon features that can increase factual coincidence, source attributions, and ethical compliance, particularly in cases of high epistemic risk.

In conclusion, the conception, implementation, and communication of scientific work represent a notable turning point. It's important to highlight not just the characteristics involved but also how the scientific community manages, adapts, and ethically integrates these elements. Thus, the key question becomes how ChatGPT will be incorporated into scientific discovery while upholding research integrity, creativity, and accountability.

6.0 REFERENCES

1. *AlZaabi, A., ALAmri, A., Albalushi, H., Aljabri, R., & AalAbdulsalam, A. (2023).* ChatGPT applications in academic research: A review of benefits, concerns, and recommendations. *bioRxiv. https://doi.org/10.1101/2023.08.17.553688*
2. *De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023).* ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. Frontiers in Public Health, 11, *1166120. https://doi.org/10.3389/fpubh.2023.1166120*
3. *Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines.* Journal of Business Research, 133, *285–296. https://doi.org/10.1016/j.jbusres.2021.04.070*
4. *Florindo, F. (2023).* ChatGPT: A threat or an opportunity for scientists? Perspectives of Earth and Space Scientists, 4*(1), e2023CN000212. https://doi.org/10.1029/2023CN000212*
5. *Homolak, J. (2023).* Opportunities and risks of ChatGPT in medicine, science, and academic publishing: A modern Promethean dilemma. Croatian Medical Journal, 64*(1), 1–3. https://doi.org/10.3325/cmj.2023.64.1*
6. *Huang, J., & Tan, M. (2023).* The role of ChatGPT in scientific communication: Writing better scientific review articles. American Journal of Cancer Research, 13*(4), 1148–1154.*
7. *Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical research ed.), 372, n71. https://doi.org/10.1136/bmj.n71*
8. *Park, Y. J., Kaplan, D., Ren, Z., Hsu, C. W., Li, C., Xu, H., et al. (2023).* Can ChatGPT be used to generate scientific hypotheses? Journal of Materiomics. *Advance online publication. https://doi.org/10.1016/j.jmat.2023.08.007*
9. *Prandner, D., Wetzelhütter, D., & Hese, S. (2024).* ChatGPT as a data analyst: An exploratory study on AI-supported quantitative data analysis in empirical research. Frontiers in Education, 9, *1417900. https://doi.org/10.3389/feduc.2024.1417900*
10. *Sharma, G., & Thakur, A. (2023).* ChatGPT in drug discovery. ChemRxiv. *https://doi.org/10.26434/chemrxiv-2023-qgs3k*
11. *Stokel-Walker, C. (2023).* ChatGPT listed as author on research papers: many scientists disapprove. Nature, 613*(7945), 620–621. https://doi.org/10.1038/d41586-023-00107-z*
12. *Teixeira da Silva, J. A. (2023).* Is ChatGPT a valid author? Nurse Education in Practice, 68, *103600. https://doi.org/10.1016/j.nepr.2023.103600*
13. *Thorp, H. H. (2023).* ChatGPT is fun, but not an author. Science, 379*(6630), 313. https://doi.org/10.1126/science.adg7879*
14. *Tong, Y., & Zhang, L. (2023).* Discovering the next decade's synthetic biology research trends with ChatGPT. Synthetic and Systems Biotechnology, 8*(2), 220–223. https://doi.org/10.1016/j.synbio.2023.02.004*
15. *Wang, K. D., Burkholder, E., Wieman, C., Salehi, S., & Haber, N. (2024).* Examining the potential and pitfalls of ChatGPT in science and engineering problem-solving. Frontiers in Education, 8, *1330486. https://doi.org/10.3389/feduc.2023.1330486*