Online Journal of Communication and Media Technologies, 2025, **15**(1), e202502
e-ISSN: 1986-3497

OPEN ACCESS

Review Article

Systematic analysis of generative AI tools integration in academic research and peer review

Husain Abdulrasool Salman [1*]
0000-0001-8241-5274

Muhammad Aliif Ahmad [2]
0000-0003-1186-531X

Roliana Ibrahim [2]
0000-0001-7580-1804

Jamilah Mahmood [2]
0000-0002-3320-5331

[1] University of Bahrain, Zallaq, BAHRAIN
[2] Universiti Teknologi Malaysia, Johor Bahru, MALAYSIA
[*] Corresponding author: habdulrasool@uob.edu.bh

Citation: Salman, H. A., Ahmad, M. A., Ibrahim, R., & Mahmood, J. (2025). Systematic analysis of generative AI tools integration in academic research and peer review. *Online Journal of Communication and Media Technologies, 15*(1), e202502. https://doi.org/10.30935/ojcmt/15832

ARTICLE INFO

Received: 19 Jun 2024
Accepted: 7 Dec 2024

ABSTRACT

While sparking a big debate among academics, generative artificial intelligence (GAI) tools are becoming integral to academic research, holding the potential to transform traditional research and peer review methods. This systematic literature review investigates the emergent role of GAI tools in academic research workflow and scholarly publications by analyzing 44 articles. The process of identifying the most relevant publications was done following the preferred reporting items for systematic reviews and meta-analyses method. The findings provide a thorough understanding of how GAI is currently being utilized in the various aspects of academic research workflow and peer review process, including concerns, limitations, and proactive measures to better employ these tools effectively. Our review suggests the need for more research to develop appropriate policies and guidelines, enhance researchers' artificial intelligence literacy through targeted training, and ensure ethical use of these tools to boost research productivity and quality.

Keywords: generative AI, chatbot, ChatGPT, academic research, scientific writing, peer review

INTRODUCTION

Quality academic research is essential for advancing knowledge, solving problems, and aiding decision-making. It introduces new perspectives and innovative solutions and helps to expand current human understanding and knowledge (Mehta, 2023). However, researchers face various challenges that can hinder their productivity and publishing efforts. Cognitive barriers, such as difficulties with complex problems requiring innovative solutions. Social barriers include lengthy journal review processes and complications in collaborative projects. Physical barriers, such as time management and data collection challenges, compounded by inadequate infrastructure, can all lead to writer's block. Additionally, the evolving nature of knowledge implies that scholarly research is continuously changing. Researchers must update their expertise and knowledge base to keep up with advancements in their field and adapt to new technologies and methodologies (Aydin et al., 2022).

Since introducing the generative pre-trained transformer AI model known as GPT 3 to the public in June 2020, it has established a standard for generative artificial intelligence (GAI) tools' capabilities. This technology

Copyright © 2025 by authors; licensee OJCMT by Bastas, CY. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).

has fundamentally transformed various domains including academic writing practices. A GAI tool is a type of "intelligent" software program that can react to human text or voice inquiries by using natural language processing and deep learning techniques (Olujimi & Ade-Ibijola, 2023). These tools make use of the capabilities offered by LLMs. LLM are machine learning models trained on datasets consisting of a large amount of data allowing them to comprehend interpret and generate responses that are contextually relevant and coherent (Hadi et al., 2023). These tools generate different types of media, such as text, voice, images, and videos, some also offer advanced tasks such as image recognition, predictive modeling, and video analysis (Nakavachara et al., 2024). Examples of popular GAI tools include ChatGPT, CoPilot, Bard, and Claude.

GAI tools are increasingly adopted in various domains. In the medical sector for instance these tools are used for tasks like training medical specialists on disease diagnosis (Zhang & Kamel Boulos, 2023). Many organizations use these tools in their web portals to help aid customer service and support, gather information, and various interactive purposes (Chen et al., 2024). In Education, they are used in different tasks such as enhancing the learning process through interaction and tailored content (Rospigliosi, 2023), assessments generation through integration with learning management systems (Pesovski et al., 2024; Salman et al., 2022), and foreign language instruction (Pérez-Núñez, 2023). These tools are also being tested to aid intelligent manufacturing (Rane et al., 2024), journalism (Pavlik, 2023), and many other fields.

GAI tools are expected to impact the traditional methods within academic research with numerous tools designed to support various tasks. Crawford et al. (2023) and Salimi and Saheb (2023) argued that these tools not only aid researchers in improving manuscript writing but also assist in literature review processes, referencing, data analysis, and peer review tasks. A study by Imran and Almusharraf (2023) reported that while only (51.5%) of academics utilize GAI in their work currently, 72.3% of them acknowledge their impact on work processes. Moreover, 83.2% of respondents plan to increase their use of these tools in the future underscoring the growing influence of GAI in the academic research domain. Furthermore, studies indicated that GAI could improve the productivity of researchers in their academic work by shifting their focus from routine tasks to higher-order analysis tasks which could lead to more innovative and impactful research outcomes (Hamamah et al., 2023; Khalifa & Albadawy, 2024). These tools could also help peer reviewers and editors work more efficiently. This support can lead to better quality reviews and help with the reviewers' shortage, suggesting productivity improvements in the peer review process (Checco et al., 2021; Hosseini & Horbach, 2023).

Despite the significant impact of GAI on academic research and peer review practices, only a few review studies were conducted in this field like Khalifa and Albadawy (2024) which reviewed studies from Google Scholar, Embase, and PubMed databases and a scoping review by Khalifa and Ibrahim (2024) based on PubMed database only. In addition, Imran and Almusharraf (2023) focused on ChatGPT as an academic writing assistant. Given the continuous evolution of these tools and their potential to transform traditional academic research practices, this review will help other researchers by conducting a systematic literature review (SLR) of studies that have examined GAI tools in academic research and peer review processes by answering the following questions:

- **RQ1:** What is the latest research progress in utilizing GAI tools for academic research and peer review processes? and what are the most researched GAI tools?
- **RQ2:** What are the primary tasks in which GAI tools are currently being tested to assess their effectiveness in academic research and peer review?
- **RQ3:** What are the main areas of concern and limitations reported by researchers regarding the use of GAI tools?
- **RQ4:** What measures would benefit the academic research community to utilize GAI tools better and overcome their limitations?

The structure of this paper is organized as follows: We first cover research methodology and research protocols. Then, we present the findings of the literature review, discussion and answers to the research questions. After that we discuss the implications of the study. Finally, we conclude the paper with final insights and limitations.

Table 1. Inclusion and exclusion criteria

Checked	Inclusion criteria	Exclusion criteria
Duplicates		Discarding duplicate studies.
Time frame	> 2020	≥ 2020
Article language	English	Other languages
Article type	Conference, Journal	Editorials, preprints, books, book chapters, notes, erratum, letters, and short essays
Scope of Article	Articles addressing the aims of the research	Non relevant articles

METHODOLOGY

To answer the research questions, this review adopts the guidelines for conducting SLRs proposed by Keele (2007). It involved three main stages: planning, conducting, and reporting. During the planning stage, we identified the need for the review, specified research questions, and developed a review protocol. In the conducting stage, we implemented a comprehensive search strategy to identify relevant studies, selected them based on predefined inclusion and exclusion criteria, assessed their quality, and extracted and synthesized the data. For the reporting stage, we determined the appropriate dissemination methods, formatted the main report, and ensured its quality through thorough evaluation. The review followed the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines (Moher et al., 2015) to systematically review literature across three recognizable databases: Web of Science, Scopus, and PubMed. The simultaneous search involved gathering data from all three databases, removing duplicates papers, and combining results.

Planning the Review

The integration of GAI tools in academic research and peer review is growing and yet to be explored. Various studies discussed potential advantages, limitations, and concerns in this regard (Victor et al., 2023). Consequently, it is crucial to offer a comprehensive overview of GAI applications in academic research and peer review. This includes exploring the current tasks in which GAI technology is experimented with, highlighting their potential advantages and disadvantages, and identifying proactive measures to overcome challenges encountered.

Inclusion and exclusion criteria

Establishing criteria for inclusion and exclusion is vital in maintaining the credibility and significance of literature reviews. This process assists researchers in choosing studies that adhere to certain criteria while eliminating irrelevant ones. This practice improves the precision and trustworthiness of the findings (Page et al., 2021). The detailed criteria of this study are depicted in **Table 1**. The time frame for the retrieved studies encompasses publications from January 2021 to March 2024.

Search procedure

The search procedure started with developing a review protocol to ensure all relevant literature was captured. Initially, the authors began with keywords to gather various relevant articles. They then read through some of these articles to find any overlooked keywords that could be crucial. The authors also consulted with field experts to provide additional insights into important terms. Since this review was intended to delve deeply into the specific area of GAI tools applications in academic research and peer review a title search strategy was adopted (Kraus et al., 2020).

Conducting the Review

Multiple initial searches were conducted in which keywords were iteratively refined based on findings, with adjustments to incorporate emerging terms. The relevant keywords and their synonyms are shown in **Table 2**. Boolean operators “OR” and “AND” in addition to the wildcard “*” were used to enhance the search for relevant studies (Mohamed Shaffril et al., 2021). **Table 2** displays the search string.

Table 2. Search string keywords

Keywords
(("ChatGPT" OR "AI bot" OR "bot" OR "LLM" OR "large language models" OR "chatbot" OR "conversational agent" OR "generative AI" OR "Bert" OR "transformers" OR "AI") AND ("peer review" OR "publication*" OR "research writing" OR "scientific writing" OR "academic writing" OR "publishing" OR "co-author" OR "researchers" OR "authorship" OR "author*" OR "academic research" OR "scholar*" OR "article" OR "literature" OR "data analysis" OR "hypothesis" OR "drafting" OR "brainstorming"))

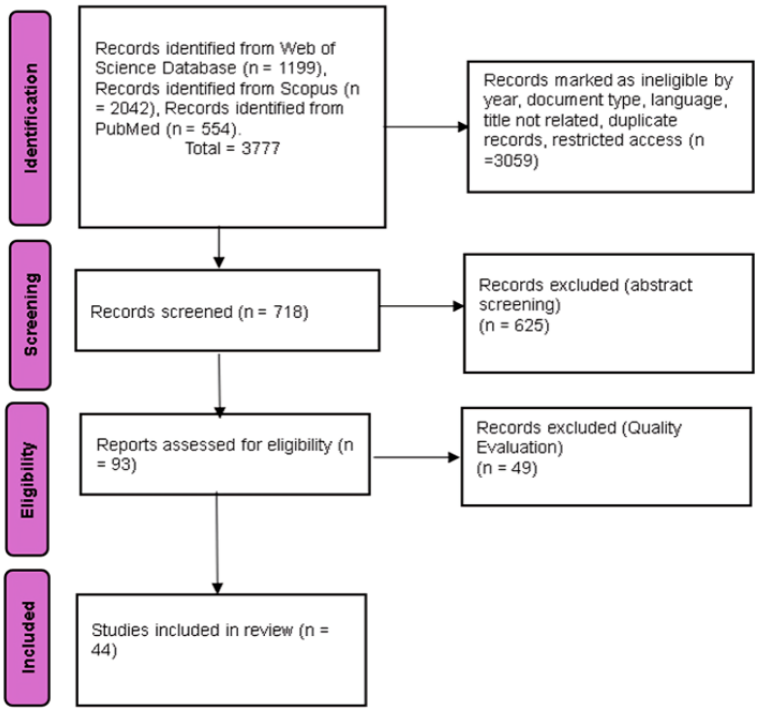

Figure 1. PRISMA flowchart (Source: Authors)

Table 3. Used quality indicators

Quality indicator (QI)	Criteria
QI-1	Alignment with inclusion and exclusion criteria
QI-2	Study design includes structured methods, clear objectives
QI-3	The article provides new knowledge
QI-4	Validity and reliability of measurement tools/theoretical integrity
QI-5	Justification of outcomes

Data extraction

The search results obtained a total of 3,777 studies which is sufficient for literature synthesis (Kraus et al., 2020). Implementing PRISMA was based on its ability to organize research articles effectively and set criteria for inclusion or exclusion. The data extraction process was performed using the inclusion and exclusion criteria in addition to the quality screening. In the eligibility stage, abstracts of 93 studies were closely examined. This resulted in the exclusion of 49 studies in the quality screening process. Ultimately, 44 articles were eligible and selected for in-depth analysis and review (**Figure 1**).

Quality evaluation

Assessing the quality of the articles chosen is a critical step (Behera et al., 2019) . Specific criteria outlined in **Table 3** were used to evaluate the quality of each article. The evaluation process involved two researchers working independently, one for data extraction and the other for verification. A rating scale of 0 to 2 was utilized for each quality indicator, a value of 0 denoting "not good" 1 denoting "good", and 2 denoting "very good." In addition, researchers' viewpoints were discussed to decide on which publications to include. A

Table 4. Quality evaluation scores

SN	Authors	Article type	QI-1	QI-2	QI-3	QI-4	QI-5	Score
1	Semrl et al. (2023)	Journal	2	2	2	2	2	10
2	Alkaissi and McFarlane (2023)	Journal	2	2	2	2	2	10
3	Praveen and Vajrobol (2023)	Journal	2	2	2	2	2	10
4	Checco et al. (2021)	Journal	2	2	2	2	2	10
5	Lozić and Štular (2023)	Journal	2	2	2	1	2	9
6	Ariyaratne et al. (2023)	Journal	2	2	2	2	2	10
7	Jarrah et al. (2023)	Journal	2	2	2	2	2	10
8	Abuyaman (2023)	Journal	2	2	2	1	2	9
9	Zohouri et al. (2024)	Journal	2	2	1	1	2	8
10	Athaluri et al. (2023)	Journal	2	2	2	2	2	10
11	Seth et al. (2023)	Journal	2	2	2	2	2	10
12	Margetts et al. (2024)	Journal	2	2	2	2	2	10
13	Livberber (2023)	Journal	2	2	2	1	2	9
14	Anghelescu et al. (2023)	Journal	2	2	1	1	1	7
15	Abdelhafiz et al. (2024)	Journal	2	2	2	1	2	9
16	Leong (2023)	Journal	2	2	1	1	1	7
17	Nazzal et al. (2024)	Journal	2	2	2	2	2	10
18	Wu and Dang (2023)	Journal	2	2	2	1	1	8
19	Nguyen et al. (2024)	Journal	2	2	2	2	2	10
20	Dashti et al. (2023)	Journal	2	2	1	1	2	8
21	Mahyoob et al. (2023)	Journal	2	2	2	1	2	9
22	Májovský et al. (2023)	Journal	2	2	2	2	2	10
23	Bond et al. (2024)	Journal	2	2	2	1	2	9
24	Zheng et al. (2023)	Journal	2	2	1	1	2	8
25	Hake et al. (2024)	Journal	2	2	1	1	2	8
26	Saad et al. (2024)	Journal	2	2	2	1	2	9
27	Perkins and Roe (2023)	Journal	2	2	2	2	2	10
28	Mollaki (2024)	Journal	2	2	2	1	1	8
29	Biswas et al. (2023)	Journal	2	2	2	1	1	8
30	Watermeyer et al. (2023)	Journal	2	2	2	2	2	10
31	Jenko et al. (2024)	Journal	2	2	1	2	2	9
32	Aiumtrakul et al. (2023)	Journal	2	2	2	2	2	10
33	Gupta et al. (2023)	Journal	2	2	2	2	2	10
34	Alshami et al. (2023)	Journal	2	2	2	2	2	10
35	Khlaif et al. (2023)	Journal	2	2	2	2	2	10
36	Dergaa et al. (2023)	Journal	2	2	2	2	2	10
37	Alyasiri et al. (2024)	Journal	2	2	2	2	2	10
38	Carabantes et al. (2023)	Journal	2	2	1	1	1	7
39	Schmidt et al. (2023)	Conference	2	2	2	1	1	8
40	Yan et al. (2023)	Conference	2	2	1	1	1	7
41	Dengel et al. (2023)	Journal	2	2	1	2	2	9
42	Burger et al. (2023)	Journal	2	2	2	2	2	10
43	Firat (2023)	Journal	2	2	1	2	2	9
44	Crawford et al. (2023)	Journal	2	2	1	1	1	7

minimum score of 7 was required for study inclusion. Of the articles reviewed, 44 surpassed this threshold and were incorporated into the study. The findings from this quality assessment are detailed in **Table 4**.

FINDINGS AND ANALYSIS

In this section, we present the outcomes derived from our search processes. Our literature search focused on journal and conference articles exploring the applications of GAI tools in academic research and peer review processes from three recognized databases. Additionally, we applied quality evaluation criteria to ensure the chosen articles were the most relevant for our study. Consequently, a total of 44 articles constituted the final dataset for this research.

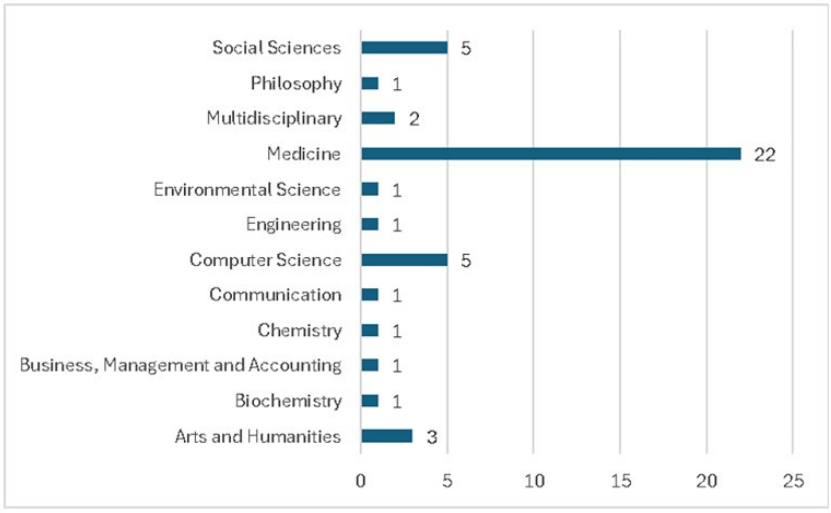

Figure 2. Fields in which GAI tools are examined for academic research and peer review (Source: Authors)

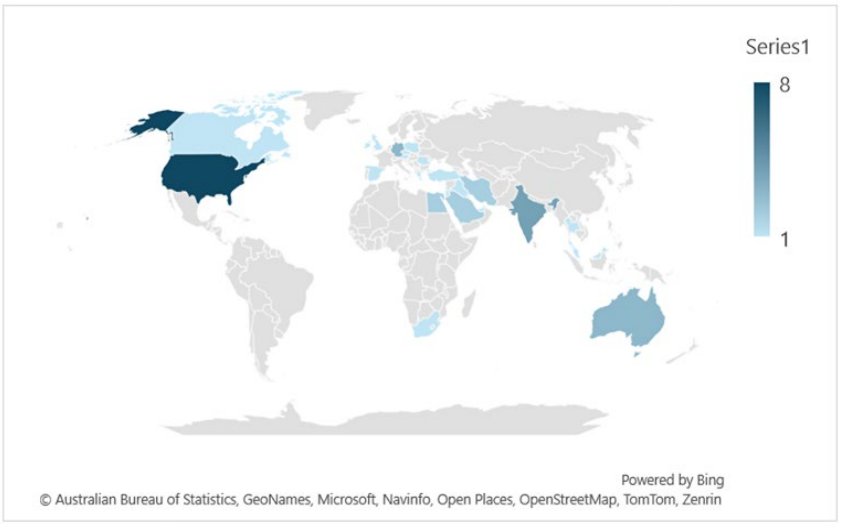

Figure 3. Geographic distribution of reviewed academic articles by country of origin of the main author (Source: Authors)

RQ1. What Is the Latest Research Progress in Utilizing GAI Tools for Academic Research and Peer Review Processes? And What Are the Most Researched Gai Tools?

The review indicates a growing number of studies examining GAI tools within academic research and peer review across diverse fields. The evolution of these tools has led to more complex capabilities and improved accuracy which is prompting researchers to explore their applications in various aspects in this area. This surge in interest underscores the significance of conducting comprehensive research in this area given the potential long-term impacts on these domains.

Figure 2 depicts a distribution of fields that examined GAI tools for academic research and peer review based on the reviewed literature. Medicine holds the highest percentage of studies with 22 studies. Computer science and social sciences follows behind constituting five studies each. Followed by arts and humanities, and multidisciplinary fields. The remaining fields, including environmental science, engineering, chemistry, and others, represent one study for each.

The data presented in **Figure 3** reveals the geographical distribution of the reviewed articles based on the main author's country of origin. The United States leads prominently with eight studies followed by India and Australia.

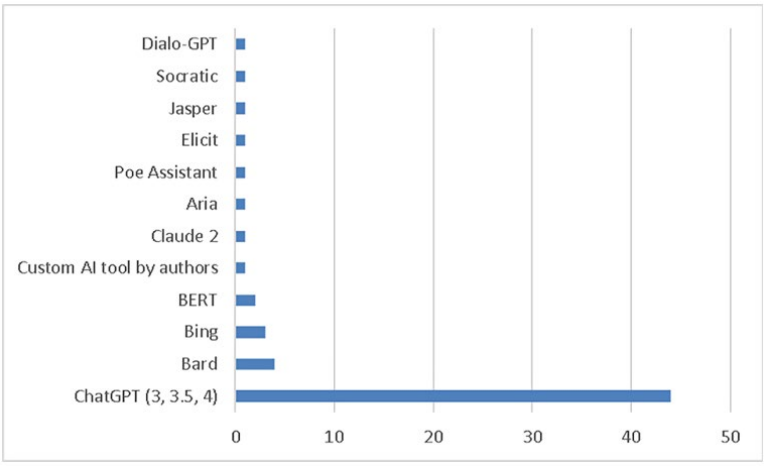

Figure 4. Researched GAI tools (Source: Authors)

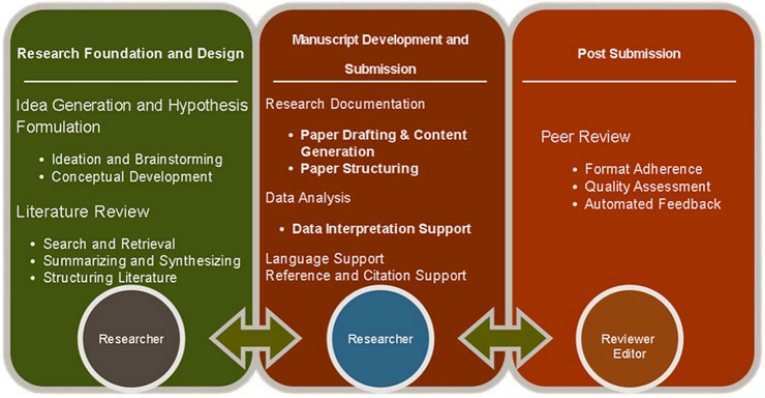

Figure 5. Current discussion of GAI tools across various stages of academic research workflow (Source: Authors)

The answer to the second part of RQ1 indicates that ChatGPT emerges as one of the most extensively tested and utilized tools by researchers as depicted in **Figure 4**. This includes its various models. In some research, there was a lack of clarity on the version of ChatGPT used. It is assumed that this pertains to the earlier editions of ChatGPT models. Studies like Dhanvijay et al. (2023) and Patil et al. (2024) have shown that ChatGPT performs better than other GAI tools in education and medicine, especially in accuracy and handling complex tasks. In addition, ChatGPT's continuous updates and extensive research support make it a preferable choice for academic use (Eigenmann et al., 2023). The review indicates few studies that have researched other GAI tools including, Bing Chat, Bard AI, Claude 2, Aria, Poe Assistant, Socratic, and Jasper. While the focus on ChatGPT is understandable given its wide reputation and performance, it is imperative that other GAI tools, especially those designed for academic research tasks receive more attention. GAI research assistant tools which are designed specifically for academic research, necessitate further exploration to determine their efficacy and potential benefits.

RQ2. What Are the Primary Tasks in Which GAI Tools Are Currently Being Tested to Assess Their Effectiveness in Academic Research and Peer Review?

Through a detailed review of the eligible studies, we have pinpointed the primary academic research tasks where GAI tools are currently being explored and tested in scholarly literature. These include "ideas generation and hypotheses formulation", "literature review", "research drafting and language support", "data analysis", "references and citation support", and "peer review". We present a breakdown of the 44 studies across six areas of academic research workflow and the peer review process. These areas are grouped into three research stages: research foundation and design, manuscript development and submission, and post-submission (as shown in **Figure 5**). The categorization into these three stages reflects the natural progression

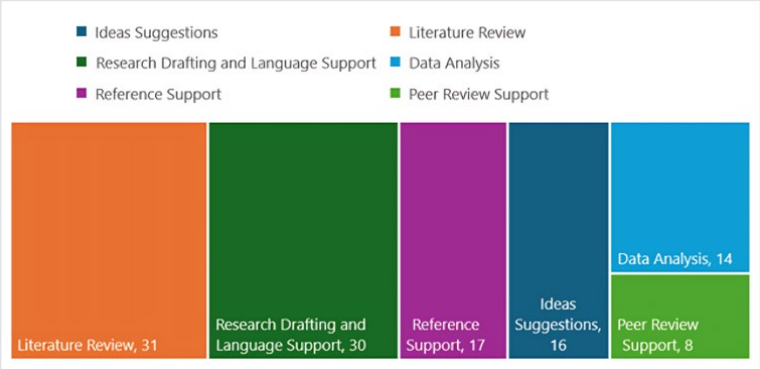

Figure 6. Distribution of GAI tool usage across academic research tasks (Source: Authors)

of a research project from conception to dissemination and organizes the key tasks of academic research and peer review that are currently under investigation for automation using GAI tools.

The review reveals a focus on the utilization of GAI tools in tasks such as "literature review" and "research drafting and language support", with these areas being the primary focus in 31 and 30 studies, respectively. This indicates a growing reliance on GAI to support research tasks like information gathering, knowledge synthesis, and manuscript documentation. Moreover, services like "reference support", "idea suggestions", and "data analysis" also indicate the ongoing evaluation of their capabilities in these areas. As shown in **Figure 6**. The academic research tasks are detailed in **Table 5** for each study.

Ideas generation and hypotheses formulation

The task of generating "ideas generation and hypotheses formulation" encompasses activities like the generation of initial research ideas, refinement of research questions, broadening conceptual scopes, and formulating initial hypotheses. Several studies discussed how GAI can assist in generating and refining research ideas. Findings indicate that these tools can aid researchers during the brainstorming phase by presenting diverse perspectives (Dashti et al., 2023; Dergaa et al., 2023; Livberber, 2023; Mahyoob et al., 2023; Mondal & Mondal, 2023). For example, Gupta et al. (2023) evaluated ChatGPT's performance to generate novel ideas for systematic review studies and found that although it sometimes produces non-accurate content, it can be utilized by researchers for this purpose effectively. Interactive engagement with these tools also helps researchers sharpen their focus (Athaluri et al., 2023; Livberber, 2023). Additionally, studies have emphasized the ability of GAI tools to help identify research gaps and emerging topics by analyzing large amounts of data (Alshami et al., 2023; Zheng et al., 2023). In formulating research hypotheses and research questions findings indicate that ChatGPT can be a supportive tool in the task, it can simplify hypothesis formulation by detecting patterns and insights within the data they are trained on (Mondal & Mondal, 2023). These studies indicated that ChatGPT can assist researchers in the early stages of research planning such as research design and methodology selection.

Literature review

One recognized advantage of GAI in the literature review stage is their ability to summarize and condense information from multiple sources (Hake et al., 2024), allowing researchers to manage large volumes of data which can be overwhelming (Datt et al., 2023; Khlaif et al., 2023; Semrl et al., 2023; Seth et al., 2023; Zohouri et al., 2024). This advantage can enhance the comprehensiveness of literature reviews (Burger et al., 2023; Dergaa et al., 2023) and effectively reduce the time spent on examining studies (Jenko et al., 2024; Margetts et al., 2024). A study by Alshami et al. (2023) demonstrated that ChatGPT-3.5 helped with high accuracy in filtering and categorizing articles, with significant time savings compared to traditional methods. Recent versions of Chat-GPT indicated better performance in this task (Abuyaman, 2023). Lozić and Štular (2023), on the other hand, tested six GAI tools and reported that while these tools help combine existing knowledge, they fail to produce original content. Other studies such as Jenko et al. (2024) reported that ChatGPT literature summaries sometimes contained factual errors and lacked critical analysis.

Table 5. Primary studies examining GAI tools for academic research and peer review

SN	Authors	IS	LR	RDLS	DA	RS	PRS	GAI Tools Examined
1	Semrl et al. (2023)		X	X		X		ChatGPT
2	Alkaissi and McFarlane (2023)		X	X		X		ChatGPT
3	Praveen and Vajrobol (2023)				X			BERT
4	Checco et al. (2021)						X	Custom AI tool by authors
5	Lozić and Štular (2023)		X	X				ChatGPT-3.5, ChatGPT-4, Bing, Bard, Claude 2, Aria
6	Ariyaratne et al. (2023)			X		X		ChatGPT-3.5
7	Jarrah et al. (2023)			X		X		ChatGPT
8	Abuyaman (2023)		X	X		X		GPT-4 and GPT-3.5
9	Zohouri et al. (2024)		X	X			X	ChatGPT
10	Athaluri et al. (2023)	X		X		X		ChatGPT
11	Seth et al. (2023)		X	X		X		ChatGPT-3
12	Margetts et al. (2024)	X	X	X		X		ChatGPT-4
13	Livberber (2023)	X	X	X		X		ChatGPT-3.5 and ChatGPT-4
14	Anghelescu et al. (2023)	X	X	X				ChatGPT-3 and GPT-3.5
15	Abdelhafiz et al. (2024)		X	X	X	X	X	ChatGPT
16	Leong (2023)		X	X	X			Bard, ChatGPT-4, Poe Assistant
17	Nazzal et al. (2024)		X	X				ChatGPT 4.0
18	Wu and Dang (2023)					X		ChatGPT
19	Nguyen et al. (2024)	X	X	X				ChatGPT-4
20	Dashti et al. (2023)	X	X	X	X	X		ChatGPT
21	Mahyoob et al. (2023)	X	X	X				ChatGPT
22	Májovský et al. (2023)			X	X			ChatGPT-3
23	Bond et al. (2024)	X	X					ChatGPT and Elicit
24	Zheng et al. (2023)	X	X	X	X			ChatGPT
25	Hake et al. (2024)		X	X				ChatGPT-3.5
26	Saad et al. (2024)		X	X		X	X	ChatGPT (versions 3.5 and 4.0)
27	Perkins and Roe (2023)	X	X	X	X	X		ChatGPT (versions 3.5 and 4.0)
28	Mollaki (2024)						X	ChatGPT
29	Biswas et al. (2023)						X	ChatGPT
30	Watermeyer et al. (2023)			X				Generative GAI
31	Jenko et al. (2024)		X					ChatGPT (GPT-4)
32	Aiumtrakul et al. (2023)		X			X		ChatGPT-3.5, Bing Chat, Bard AI
33	Gupta et al. (2023)	X	X					ChatGPT
34	Alshami et al. (2023)	X	X					ChatGPT-3.5
35	Khlaif et al. (2023)		X	X		X		ChatGPT
36	Dergaa et al. (2023)	X	X		X			ChatGPT
37	Alyasiri et al. (2024)		X	X		X		ChatGPT-4
38	Carabantes et al. (2023)				X		X	ChatGPT (versions 3.5 and 4)
39	Schmidt et al. (2023)	X						ChatGPT (versions 3.5 and 4)
40	Yan et al. (2023)		X		X			ChatGPT-4
41	Dengel et al. (2023)	X		X	X			ChatGPT-4, Bert
42	Burger et al. (2023)		X	X	X			ChatGPT-3
43	Firat (2023)	X	X	X	X			GPT-4 and GPT-3.5
44	Crawford et al. (2023)		X	X	X		X	ChatGPT-3, Bard, Bing, Jasper, Socratic, Dialo-GPT

Note. IS: Ideas suggestions; LR: Literature review; RDLS: Research drafting and language support; DA: Data analysis; RS: Reference support; PRS: Peer review support.

In addition, Crawford et al. (2023), Margetts et al. (2024), and Nazzal et al. (2024) compared the time and accuracy of writing literature review articles using three methods, AI-only, AI-assisted, and human-only, and found that while the AI-only method took the least time to complete, it had the lowest accuracy and required substantial time for revision. Recent AI literature review assistants such as "AI Scholar" and "Elicit" can be used to overcome previous limitations. For instance, a study by Bond et al. (2024) reported that combining ChatGPT and Elicit to conduct a literature review can streamline the review process by quickly identifying and examining specific sections of relevant studies.

Research drafting and language support

The basic capabilities of GAI show that these tools efficiently assist in the preliminary stages of writing, for instance, they enable researchers to produce initial drafts with greater speed compared to the traditional writing process (Abuyaman, 2023; Alkaissi & McFarlane, 2023; Firat, 2023; Leong, 2023; Lozić & Štular, 2023;

Mahyoob et al., 2023; Semrl et al., 2023; Zheng et al., 2023). This allows researchers to dedicate more time to critical thinking and analysis rather than focusing on writing techniques. In addition, Ariyaratne et al. (2023) and Margetts et al. (2024) explored how GAI technologies are continuously evolving and provide more services to researchers, such as helping to draft complete parts of manuscripts like introductions and conclusions. This suggests a shift from simple drafting tools to more comprehensive writing assistance. Furthermore, several studies reported services like enhancing existing texts in terms of writing style, grammar, and overall text coherence which improves the readability and outreach of scientific work to a wider audience (Anghelescu et al., 2023; Athaluri et al., 2023; Bond et al., 2024; Dashti et al., 2023; Livberber, 2023; Margetts et al., 2024; Nguyen et al., 2024). Bekker (2024) discussed the advantages of GAI for non-native English scholars by providing translation services as an added benefit.

However, with this expanded role in manuscript documentation comes the need for careful human oversight (Nazzal et al., 2024; Perkins & Roe, 2023) since these tools can provide fabricated content (Májovský et al., 2023) and inaccurate scientific content based on biased data (Mahyoob et al., 2023). In other words, while GAI can efficiently produce complete drafts, these documents require substantial human input to ensure they meet the rigorous standards of accuracy and authenticity expected in scholarly work. In addition, users who actively interact with GAI tools achieve better academic writing (Nguyen et al., 2024). Khlaif et al. (2023) emphasized the importance of detailed "prompt engineering" to enhance the factual accuracy of the generated output.

Data analysis

Several studies in different domains indicated that GAI tools are increasingly utilized in various aspects of data analysis and interpretation, enabling researchers to perform tasks such as statistical data analysis, and data visualization, in addition to thematic analysis, ChatGPT was found to simplify the process for novice researchers to comprehend and manipulate datasets through a user-friendly analysis tool (Abdelhafiz et al., 2024; Burger et al., 2023; Crawford et al., 2023; Datt et al., 2023; Firat, 2023; Morocco-Clarke et al., 2023; Praveen & Vajrobol, 2023). In addition, GAI tools can be used by researchers to create sample datasets and assist in text classification tasks (Ebert & Louridas, 2023). In qualitative research, GAI tools enhance analysis by streamlining coding processes and assisting with the initial exploration of datasets (Dengel et al., 2023; Yan et al., 2023). Some researchers tested ChatGPT's potential in predictive modeling. For example, a study by Zheng et al. (2023) reported ChatGPT's ability to predict how certain chemicals are made by synthesizing data from 228 papers.

However, other studies reported the necessity to review generated content due to the possibility of generating inaccurate output (Semrl et al., 2023), for example, Biswas et al. (2023) tested the capability of using ChatGPT to analyze medical case reports and found that it lacked synthesizing complex cases. Additionally, in their 2023 study, Zheng et al. (2023) highlighted two main challenges: the occurrence of "hallucinations" where the AI generates fabricated information, and the need for effective prompt engineering to guide the AI's responses accurately.

Reference support

In evaluating the effectiveness of GAI tools such as ChatGPT, Bard, and Bing in References and citation support, various studies have reported limitations in this task. As reported by multiple studies these tools, while beneficial in enhancing manuscript documentation and language assistance, often fall short of providing accurate references. Many studies reported inaccurate reference generation by these tools (Anghelescu et al., 2023; Ariyaratne et al., 2023; Athaluri et al., 2023; Livberber, 2023; Margetts et al., 2024; Seth et al., 2023; Wu & Dang, 2023). For example, a study by Aiumtrakul et al. (2023) in medical research evaluated the accuracy of references generated from ChatGPT-3.5, Bing Chat, and Bard in different nephrology topics and revealed that the references generated included a mix of correct and incorrect, with some being incomplete or entirely fabricated. These studies emphasize the importance of a detailed review by researchers to ensure the accuracy of AI-generated references and citations. However, the latest GAI research plug-ins such as "Consensus" and "Scholar GPT" which are based on the GPT model do provide access to databases such as Google Scholar, PubMed, JSTOR, etc.

Table 6. Concerns and limitations discussed on GAI use

Concern/limitation	Details
1. Accuracy and bias	
– Data erroneous	Factual inaccuracies, incorrect or non-relevant content
– Outdated data	Non-recent training data
– Fabricated data	Non-existing, made-up content
– Biased data	Skewed or unfair content
2. Academic integrity	
– Plagiarism and originality	Potential for AI to produce plagiarized or insufficiently original content
– Ethical standards and verification	Lack of robust ethical guidelines and verification processes to ensure the originality and integrity of AI-generated content
3. Reliance on GAI	
– Impact on research skills	Reduced development of critical research skills and lower research quality due to AI dependency
– Impact on innovation and creativity	Decreased innovation and creative thinking as reliance on GAI tools
– Increased workload	Expectations of more work from academics given the existence of GAI tools
4. Data privacy	
– Data misuse by GAI tools providers	Risks of data being used unethically, including privacy breaches and misuse of personal data
– Vulnerability to cyber-attacks	Data breaches, and unauthorized access by cyber-attackers

Peer review

The discussion of integrating GAI in the peer review process is growing across the academic community. Various studies discussed the role of ChatGPT that could play in streamlining this critical process (Biswas et al., 2023; Carabantes et al., 2023; Checco et al., 2021; Saad et al., 2024). Findings indicate that it can perform tasks like automating initial screenings and assessing the quality of manuscripts which could significantly lighten the load on reviewers. Additionally, there is potential to provide feedback on manuscripts (Checco et al., 2021) and maintain structural and content quality with academic standards (Abdelhafiz et al., 2024; Zohouri et al., 2024). Despite these advantages, the current technical limitations of the GAI tools cannot replace human reviewers due to the lack of understanding of complex contexts (Carabantes et al., 2023; Zohouri et al., 2024). In addition, the practice of using these tools by reviewers underscores the necessity for clear guidelines and policies, as argued by Biswas et al. (2023) and Mollaki (2024).

Moreover, with AI's growing presence in peer review, establishing a framework for its use is crucial to ensure that it augments rather than undermines the integrity of the scholarly review process. Furthermore, concerns about bias and the ability of reviewers to distinguish between AI and human-generated texts require developing effective AI detection tools that can be adopted in the peer review process (Checco et al., 2021).

RQ3. What Are the Main Areas of Concern and Limitations Reported by Researchers Regarding the Use of GAI Tools?

This study also identified the main areas of researchers' concerns and GAI tools limitations associated with their use in academic research and peer review and classified them into four categories (**Table 6**). "Accuracy and bias", this aspect includes reported issues related to the accuracy of data generated and subdivided into erroneous, outdated, fabricated, and biased data. The second concern, "academic integrity and ethical use" relates to concerns related to the ethical use of GAI tools like plagiarism and originality, transparent use, and privacy of data.

Concerns related to the impact on researchers' skills and research innovation were grouped under the "dependence on AI" theme. The fourth aspect discussed by researchers was the "data privacy and security" of GAI tools and subdivided into, handling of complex data and context window size.

Accuracy and bias

Findings indicate that the "accuracy and bias" of the GAI-generated content is a significant concern among researchers. This aspect is related to data erroneous like factual inaccuracies and irrelevant content which is a highly threatening factor to academic research standards (Abdelhafiz et al., 2024; Abuyaman, 2023; Alkaissi & McFarlane, 2023; Bond et al., 2024). It was also reported that GAI tools can provide "outdated data" due to

access restrictions to recent scholarly work and digital databases (Abuyaman, 2023; Májovský et al., 2023; Mohammed Alyasiri et al., 2024). Other authors reported fabricated non-existent generated content by GAI tools known as "AI Hallucinations" which is a significant challenge in GAI development because it undermines their reliability and trustworthiness (Alkaissi & McFarlane, 2023; Carabantes et al., 2023; Firat, 2023; Mahyoob et al., 2023; Schmidt et al., 2023; Zheng et al., 2023).

Another concern raised is that GAI-generated content can be biased. This issue can occur if the training data lacks diverse perspectives which can lead to biased output threatening academic inclusivity (Abuyaman, 2023; Hake et al., 2024; Mollaki, 2024; Perkins & Roe, 2023). Generated content Inaccuracies occur due to technical limitations of GAI Tools. These studies highlighted issues where AI tools struggle to process and interpret complex datasets or non-standard formats, limiting their effectiveness in fields requiring deep contextual understanding. Additionally, their limitation to processing large amounts of data due to the context window size affects their efficiency in logical reasoning tasks as discussed by Carabantes et al. (2023). Furthermore, GAI models, which rely on deep neural networks, lack explainability because their decision-making processes are complex and difficult to interpret.

Academic integrity

Another significant concern is the potential for GAI tools to compromise academic integrity. Numerous studies have underscored significant issues concerning plagiarized academic content, whether it is done intentionally or unintentionally by researchers. This could happen if the GAI output closely resembles existing scientific material without proper acknowledgment or attribution (Crawford et al., 2023; Jarrah et al., 2023; Livberber, 2023; Margetts et al., 2024; Zohouri et al., 2024). These studies suggested the formation of policies and guidelines to govern and promote responsible and ethical use of GAI tools in academic research workflow and peer review processes to maintain the integrity of the scholarly work and the development of robust AI-text detection models to help identify plagiarized text effectively (Abdelhafiz et al., 2024; Mollaki, 2024; Perkins & Roe, 2023).

Reliance on GAI

Over-reliance on GAI tools in academic research raised concerns about its impact on research skills and creativity. Some studies suggest that while GAI can enhance efficiency, it may diminish critical thinking and innovative capacities among researchers by providing easy solutions to complex problems (Checco et al., 2021; Livberber, 2023; Nguyen et al., 2024; Praveen & Vajrobol, 2023). Nguyen et al. (2024) and Zohouri et al. (2024) further warns of the potential for GAI to lead researchers to depend too heavily on automated processes, which could negatively affect genuine scholarly inquiry and innovation. In addition, as GAI advances, many researchers feel diminished by its capabilities which raises worries regarding the redundancy of the workforce. This anxiety comes from the fear that GAI tools might exceed human creativity which leads to psychological stress and potentially hinders human innovation leading to what is called "creative displacement anxiety". This concern can reduce researchers' acceptance of GAI technology and lower adoption levels. Moreover, while GAI tools can handle repetitive tasks, freeing up time for academics, this time savings might lead institutions to expect more output including academic research (Watermeyer et al., 2023).

Data privacy

Many studies raised concerns about the security risks to data privacy. Providers of GAI services could jeopardize users' privacy by mishandling their data. This danger is especially evident when using GAI tools may lead to data breaches that expose sensitive information. In addition, GAI systems are increasingly vulnerable to cyber-attacks. These systems can be targeted in different ways such as introducing harmful data or exploiting system weaknesses to gain unauthorized access. The ethical and security considerations surrounding GAI tools are significant in academic research and peer review processes and need to be addressed to ensure the confidentiality of researchers' scholarly work (Alshami et al., 2023; Anghelescu et al., 2023; Bond et al., 2024; Mollaki, 2024; Nguyen et al., 2024).

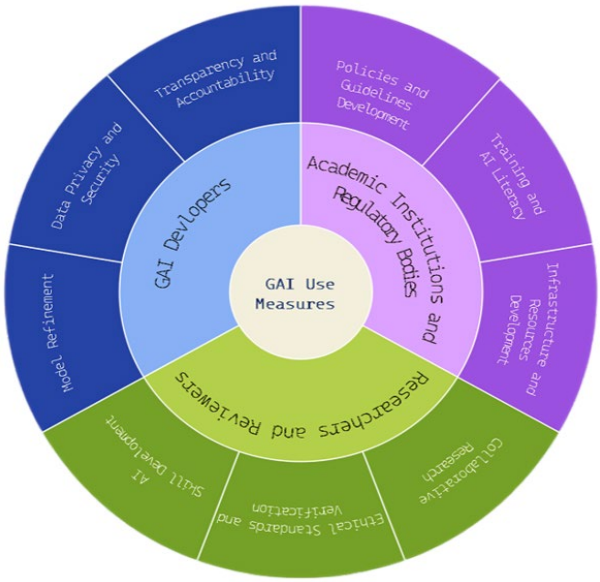

Figure 7. Measures for effective and ethical GAI use (Source: Authors)

RQ4. What Measures Would Benefit the Academic Research Community to Utilize GAI Tools Better and Overcome Their Limitations?

From the studies examined, several key measures have been identified to ensure the effective and ethical use of GAI tools. **Figure 7** illustrates a comprehensive view of these measures to enhance the use of generative GAI tools within the academic research community. It is organized into three concentric circles representing key stakeholder groups: regulatory bodies and academic institutions, researchers and reviewers, and GAI tools providers. Each group is assigned specific measures that address critical concern areas. For regulatory bodies and academic institutions, the measures include policies and guidelines development, training and AI literacy, and infrastructure and resources development. Researchers and reviewers are focused on ethical standards and verification, collaborative research as well as skill development and innovation. GAI tools providers are responsible for transparency and accountability in their development practices, data privacy and security, and model refinement. These stakeholders must work with each other not in isolation to achieve effective outcomes and enhance utilization levels while minimizing risks and concerns.

Academic institutions and regulatory bodies

Policy and regulations development: Academic institutions and regulatory bodies such as higher education, research institutions, publishers, and related government agencies need to establish clear policies to regulate the use of GAI tools in research. These policies should comprehensively address roles and responsibilities, ethical use, and guide researchers on the appropriate use of GAI. Key components of these policies should include verification procedures guidelines and requirements for transparency and accountability (Abdelhafiz et al., 2024; Alkaissi & McFarlane, 2023; Jarrah et al., 2023; Mollaki, 2024). Moreover, these policies should be reviewed and updated regularly to reflect continuous advancements in GAI technology (Livberber, 2023; Perkins & Roe, 2023; Praveen & Vajrobol, 2023).

Training and AI literacy: The second measure emphasizes the importance of providing training and enhancing researchers' AI literacy to equip them with the necessary skills to better utilize GAI tools. This could include AI skills development programs and ethical research workshops. Studies emphasized the need for training that focuses on ethical use (Abdelhafiz et al., 2024; Al-Zahrani, 2023). Nguyen et al. (2024) highlighted the importance of training strategies that meet the specific needs of researchers. For instance, some studies discussed the benefits of prompt engineering training to generate the most accurate and relevant responses which improves the precision and reliability of GAI outputs (Khlaif et al., 2023; Mahyoob et al., 2023; Zheng et al., 2023).

Resource support: Findings highlight the significance of academic institutions and government entities dedicating resources to advancing and incorporating AI tools into academic research. This involves allocating funds and infrastructure support for research programs by establishing AI research hubs. Academic institutions are also required to promote collaborative research from different disciplines to foster ethical and effective use (Al-Zahrani, 2023; Jarrah et al., 2023; Mollaki, 2024; Perkins & Roe, 2023). Additionally, providing AI-assisted research tools including GAI detection tools for researchers in academic can help enhance overall the quality and productivity of academic research (Jarrah et al., 2023; Mollaki, 2024; Perkins & Roe, 2023).

GAI providers

Transparency and accountability: GAI providers must prioritize both transparency and accountability. Transparency involves clearly explaining how AI models are developed and function. This practice will foster trust among researchers and reviewers toward the integration of these tools into academic research and peer review. In addition, AI companies must establish systems to monitor the use of their GAI tools to ensure that they are used ethically. This includes setting up measures to detect and prevent unethical practices like plagiarism and acting if misuse occurs. By holding themselves accountable, GAI providers ensure their tools support responsible usage (Yan et al., 2023).

Data privacy and security: Providers should follow ethical standards when handling sensitive data of GAI users. This includes enforcing data usage policies and robust security protocols for data protection through methods such as advanced encryption, regular security checks, and compliance with data protection laws (Anghelescu et al., 2023; Jarrah et al., 2023).

AI model refinement: GAI providers must conduct refinement audits of their models because this process addresses biases and enhances the accuracy and reliability of AI-generated outputs. Refinement audits involve evaluating both the training data and algorithm performance which helps in detecting and mitigating biases that may be embedded within the data or the algorithms themselves (Abuyaman, 2023; Alkaissi & McFarlane, 2023; Anghelescu et al., 2023; Biswas et al., 2023; Zohouri et al., 2024).

Researchers and reviewers

Collaborative research: Conducting collaborative research using GAI tools can make information more accessible and integrate diverse perspectives. This practice leads to more innovative and comprehensive research outcomes. In addition, these tools facilitate global collaboration, allowing researchers from different disciplines and locations to contribute to and build upon each other's work thereby accelerating scientific progress and promoting responsible and ethical use (Abuyaman, 2023; Alkaissi & McFarlane, 2023; Bond et al., 2024; Dergaa et al., 2023; Jarrah et al., 2023; Mollaki, 2024; Semrl et al., 2023; Zohouri et al., 2024).

Ethical standards and verification: Researchers are required to clearly state the extent and nature of GAI involvement in their studies. This ensures transparency and academic integrity which helps maintain the accuracy and reliability of academic research. In addition, adhering to ethical research standards involves verifying AI-generated content for accuracy and potential biases in GAI output (Alkaissi & McFarlane, 2023; Praveen & Vajrobol, 2023). In the review process, while GAI tools offer benefits, they are not ready to replace reviewers due to their limited capabilities in this task. Reviewers must use these tools as supplementary aids and verify their output for potential inaccuracies to ensure the integrity of the review process (Biswas et al., 2023; Checco et al., 2021; Saad et al., 2024).

AI skill development: While academic institutions are requested to provide training to enhance researchers' AI skills, it is also essential for researchers and reviewers to proactively develop these skills on their own. Researchers and reviewers need to enhance their AI literacy and prompt engineering skills to effectively utilize these tools in academic research. This is crucial because understanding the limitations of GAI and the importance of precise prompt input can significantly improve the quality and reliability of the output (Abdelhafiz et al., 2024; Khlaif et al., 2023; Zheng et al., 2023).

IMPLICATIONS

This review contributes to the knowledge of GAI tools' integration into academic research workflow and peer review. The review is one of very few current works in this specific area of research. It provides a

comprehensive view of academic research workflow tasks that are being automated using GAI tools including the reviewing process. In addition, detailed and structured concerns and limitations were uncovered. Moreover, measures to use GAI tools effectively and ethically were analyzed and organized. These findings will allow for a broader understanding and provide further directions for future research. In addition, the findings will help various stakeholders in different sectors including researchers, reviewers, policymakers, and developers take into consideration their roles and responsibilities to overcome possible drawbacks and enhance academic research productivity and quality.

CONCLUSION

This study presented systematic literature on applications of GAI tools in various tasks of academic research workflow and peer review process. The study followed the PRISMA protocol to examine 44 journal and conference articles from three databases, "Scopus", "Web of Science", and "PubMed" The findings described the academic research and peer review tasks that are currently experimented with GAI tools, including potential benefits, limitations, and concerns posed by researchers' experiments. In addition, this study identified measures and recommendations to use these tools better and overcome their limitations.

Limitations

This review exhibits a few limitations. Primarily, due to the evolving nature of GAI tools some literature, although published recently might not encompass all the new features of this technology. Additionally, the study focused only on conference and journal articles from three databases, other databases might provide more studies targeting other stages of Academic research in which GAI tools are being examined. Moreover, the study concentrated on Generative text AI tools and did not consider other forms of AI applications that can be utilized in academic research and peer review processes.

Future Research

With the release of enhanced versions of GAI tools, there is growing potential to significantly impact how they might shape the creation, revision, and dissemination of academic content. Future investigations are required to explore GAI tools' applications, particularly AI research assistants, in different stages of research and peer review and different disciplines.

Additionally, the debate over integrating GAI tools necessitates further research to explore researchers' perceptions and attitudes toward this technology. This research should measure adoption levels within specific demographics, both at individual and organizational levels. Such insights will inform the development of tailored policies that address the unique requirements and concerns of each academic institution.

Author contributions: HAS, MABA, RBI, & JBM: conception and design; **HAS & MABA:** material preparation and data collection and analysis; **HAS:** writing the first draft. All authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: This study does not contain any studies with human or animal subjects performed by any of the authors; thus, it does not require any ethical approval.

Declaration of interest: The authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Abdelhafiz, A. S., Ali, A., Maaly, A. M., Ziady, H. H., Sultan, E. A., & Mahgoub, M. A. (2024). Knowledge, perceptions and attitude of researchers towards using ChatGPT in research. *Journal of Medical Systems, 48*, Article 26. https://doi.org/10.1007/s10916-024-02044-4

Abuyaman, O. (2023). Strengths and weaknesses of ChatGPT models for scientific writing about medical vitamin B12: Mixed methods study. *JMIR Formative Research, 7*, e49459. https://doi.org/10.2196/49459

Aiumtrakul, N., Thongprayoon, C., Suppadungsuk, S., Krisanapan, P., Miao, J., Qureshi, F., & Cheungpasitporn, W. (2023). Navigating the landscape of personalized medicine: The relevance of ChatGPT, BingChat, and Bard AI in nephrology literature searches. *Journal of Personalized Medicine, 13*(10), Article 1457. https://doi.org/10.3390/jpm13101457

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus, 15*(2), Article e35179. https://doi.org/10.7759/cureus.35179

Alshami, A., Elsayed, M., Ali, E., Eltoukhy, A. E., & Zayed, T. (2023). Harnessing the power of ChatGPT for automating systematic review process: Methodology, case study, limitations, and future directions. *Systems, 11*, Article 351. https://doi.org/10.3390/systems11070351

Alyasiri, O. M., Salman, A. M., & Salisu, S. (2024). ChatGPT revisited: Using ChatGPT-4 for finding references and editing language in medical scientific articles. *Journal of Stomatology, Oral and Maxillofacial Surgery, 125*(5), Article 101842. https://doi.org/10.1016/j.jormas.2024.101842

Anghelescu, A., Ciobanu, I., Munteanu, C., Anghelescu, L. A. M., & Onose, G. (2023). ChatGPT: "To be or not to be" … in academic research. the human mind's analytical rigor and capacity to discriminate between ai bots' truths and hallucinations. *Balneo and PRM Research Journal, 14*, Article 4. https://doi.org/10.12680/balneo.2023.614

Ariyaratne, S., Iyengar, K. P., Nischal, N., Chitti Babu, N., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology, 52*, 1755–1758. https://doi.org/10.1007/s00256-023-04340-5

Athaluri, S. A., Manthena, S. V., Kesapragada, V. S. R. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: Investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus, 15*(4), Article e37432. https://doi.org/10.7759/cureus.37432

Aydin, A., Yürük, S. E., Reisoğlu, İ., & Goktas, Y. (2022). Main barriers and possible enablers of academicians while publishing. *Scientometrics, 128*, 623–650. https://doi.org/10.1007/s11192-022-04528-x

Behera, R. K., Bala, P. K., & Dhir, A. (2019). The emerging role of cognitive computing in healthcare: A systematic literature review. *International Journal of Medical Informatics, 129*, 154–166. https://doi.org/10.1016/j.ijmedinf.2019.04.024

Bekker, M. (2024). Large language models and academic writing: Five tiers of engagement. *South African Journal of Science, 120*(1/2). https://doi.org/10.17159/sajs.2024/17147

Biswas, S., Dobaria, D., & Cohen, H. L. (2023). ChatGPT and the future of journal reviews: A feasibility study. *Yale Journal of Biology and Medicine, 96*, 415–420. https://doi.org/10.59249/SKDH9286

Bond, A., Cilliers, D., Retief, F., Alberts, R., Roos, C., & Moolman, J. (2024). Using an artificial intelligence chatbot to critically review the scientific literature on the use of artificial intelligence in environmental impact assessment. *Impact Assessment and Project Appraisal, 42*(2), 189–199. https://doi.org/10.1080/14615517.2024.2320591

Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. (2023). On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management, 26*, 233–241. https://doi.org/10.1108/EJIM-02-2023-0156

Carabantes, D., González-Geraldo, J. L., & Jover, G. (2023). ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews. *El Profesional de la Informacion, 32*(5). https://doi.org/10.3145/epi.2023.sep.16

Checco, A., Bracciale, L., Loreti, P., Pinfield, S., & Bianchi, G. (2021). AI-assisted peer review. *Humanities and Social Sciences Communications, 8*, Article 25. https://doi.org/10.1057/s41599-020-00703-8

Chen, T., Gascó-Hernandez, M., & Esteve, M. (2024). The adoption and implementation of artificial intelligence chatbots in public organizations: Evidence from US state governments. *The American Review of Public Administration, 54*, 255–270. https://doi.org/10.1177/02750740231200522

Crawford, J., Cowling, M., Ashton-Hay, S., Kelder, J.-A., Middleton, R., & Wilson, G. (2023). Artificial intelligence and authorship editor policy: ChatGPT, Bard Bing AI, and beyond. *Journal of University Teaching and Learning Practice, 20*(5). https://doi.org/10.53761/1.20.5.01

Dashti, M., Londono, J., Ghasemi, S., & Moghaddasi, N. (2023). How much can we rely on artificial intelligence chatbots such as the ChatGPT software program to assist with scientific writing? *The Journal of Prosthetic Dentistry*. https://doi.org/10.1016/j.prosdent.2023.05.023

Datt, M., Sharma, H., Aggarwal, N., & Sharma, S. (2023). Role of ChatGPT-4 for medical researchers. *Annals of Biomedical Engineering, 52*, 1534–1536. https://doi.org/10.1007/s10439-023-03336-5

Dengel, A., Gehrlein, R., Fernes, D., Görlich, S., Maurer, J., Pham, H. H., Großmann, G., & Eisermann, N. D. G. (2023). Qualitative research methods for large language models: Conducting semi-structured interviews with ChatGPT and BARD on computer science education. *Informatics, 10*(4), Article 78. https://doi.org/10.3390/informatics10040078

Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport, 40*, 615–622. https://doi.org/10.5114/biolsport.2023.125623

Dhanvijay, A. K. D., Pinjar, M. J., Dhokane, N., Sorte, S. R., Kumari, A., & Mondal, H. (2023). Performance of large language models (ChatGPT, Bing Search, and Google Bard) in solving case vignettes in physiology. *Cureus, 15*(8), Article e42972. https://doi.org/10.7759/cureus.42972

Ebert, C., & Louridas, P. (2023). Generative AI for software practitioners. *IEEE Software, 40*, 30–38. https://doi.org/10.1109/MS.2023.3265877

Eigenmann, P., Akenroye, A., Markovic, M. A., Candotti, F., Ebisawa, M., Genuneit, J., Kalayci, Ö., Kollmann, D., Leung, A. S. Y., & Peters, R. L. (2023). Pediatric allergy and immunology (PAI) is for polishing with artificial intelligence, but careful use. *Pediatric Allergy and Immunology, 34*(9), Article e14023. https://doi.org/10.1111/pai.14023

Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning & Teaching, 6*(1), 1–22. https://doi.org/10.37074/jalt.2023.6.1.22

Gupta, R., Park, J. B., Bisht, C., Herzog, I., Weisberger, J., Chao, J., Chaiyasate, K., & Lee, E. S. (2023). Expanding cosmetic plastic surgery research with ChatGPT. *Aesthetic Surgery Journal, 3*, 930–937. https://doi.org/10.1093/asj/sjad069

Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage. *TechRxiv*. https://doi.org/10.36227/techrxiv.23589741.v1

Hake, J., Crowley, M., Coy, A., Shanks, D., Eoff, A., Kirmer-Voss, K., Dhanda, G., & Parente, D. J. (2024). Quality, accuracy, and bias in ChatGPT-based summarization of medical abstracts. *Annals of Family Medicine, 22*, 113–120. https://doi.org/10.1370/afm.3075

Hamamah, Emaliana, I., Hapsari, Y., Degeng, P. D. D., & Fadillah, A. C. (2023). Using nominal group technique to explore publication challenges and the usefulness of AI-based writing technologies: Insights from Indonesian scholars. *Theory and Practice in Language Studies, 13*(8), 2038–2047. https://doi.org/10.17507/tpls.1308.20

Hosseini, M., & Horbach, S. P. J. M. (2023). Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Research Integrity and Peer Review, 8*, Article 4. https://doi.org/10.1186/s41073-023-00133-5

Huang, J., & Tan, M. (2023). The role of ChatGPT in scientific communication: Writing better scientific review articles. *American Journal of Cancer Research, 13*, Article 1148.

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4), Article ep464. https://doi.org/10.30935/cedtech/13605

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say? *Online Journal of Communication and Media Technologies, 13*(4), Article e202346. https://doi.org/10.30935/ojcmt/13572

Jenko, N., Ariyaratne, S., Jeys, L., Evans, S., Iyengar, K., & Botchu, R. (2024). An evaluation of AI generated literature reviews in musculoskeletal radiology. *The Surgeon, 22*(3), 194–197. https://doi.org/10.1016/j.surge.2023.12.005

Keele, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. EBSE.

Khalifa, A. A., & Ibrahim, M. A. (2024). Artificial intelligence (AI) and ChatGPT involvement in scientific and medical writing, a new concern for researchers. A scoping review. *Arab Gulf Journal of Scientific Research, 42*(4), 1770–1787. https://doi.org/10.1108/AGJSR-09-2023-0423

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*, Article 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Khlaif, Z. N., Mousa, A., Hattab, M. K., Itmazi, J., Hassan, A. A., Sanmugam, M., & Ayyoub, A. (2023). The potential and concerns of using AI in scientific research: ChatGPT performance evaluation. *JMIR Medical Education, 9*, Article e47049. https://doi.org/10.2196/47049

Kraus, S., Breier, M., & Dasí-Rodríguez, S. (2020). The art of crafting a systematic literature review in entrepreneurship research. *International Entrepreneurship and Management Journal, 16*, 1023–1042. https://doi.org/10.1007/s11365-020-00635-4

Leong, A. P. (2023). Clause complexing in research-article abstracts: Comparing human- and AI-generated texts. *ExELL, 11*(2), 9–132. https://doi.org/10.2478/exell-2023-0008

Livberber, T. (2023). Toward non-human-centered design: Designing an academic article with ChatGPT. *Profesional de la Informacion, 32*(5). https://doi.org/10.3145/epi.2023.sep.12

Lozić, E., & Štular, B. (2023). Fluent but not factual: A comparative analysis of ChatGPT and other AI chatbots' proficiency and originality in scientific writing for humanities. *Future Internet, 15*(10), Article 336. https://doi.org/10.3390/fi15100336

Mahyoob, M., Algaraady, J., & Alblwi, A. (2023). A proposed framework for human-like language processing of ChatGPT in academic writing. *International Journal of Emerging Technologies in Learning, 18*(14), 282–293. https://doi.org/10.3991/ijet.v18i14.41725

Májovský, M., Černý, M., Kasal, M., Komarc, M., & Netuka, D. (2023). Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *Journal of Medical Internet Research, 25*, Article e46924. https://doi.org/10.2196/46924

Margetts, T. J., Karnik, S. J., Wang, H. S., Plotkin, L. I., Oblak, A. L., Fehrenbacher, J. C., Kacena, M. A., & Movila, A. (2024). Use of AI language engine ChatGPT 4.0 to write a scientific review article examining the intersection of Alzheimer's disease and bone. *Current Osteoporosis Reports, 22*, 177–181. https://doi.org/10.1007/s11914-023-00853-z

Mehta, J. (2023). Growing importance of academic research in education. *EPRA International Journal of Research & Development, 8*, 261–264.

Mohamed Shaffril, H. A., Samsuddin, S. F., & Abu Samah, A. (2021). The ABC of systematic literature review: The basic methodological guidance for beginners. *Quality & Quantity, 55*, 1319–1346. https://doi.org/10.1007/s11135-020-01059-6

Mohammed Alyasiri, O., Salman, A. M., Akhtom, D., & Salisu, S. (2024). ChatGPT revisited: Using ChatGPT-4 for finding references and editing language in medical scientific articles. *Journal of Stomatology, Oral and Maxillofacial Surgery, 125*(5, Supplement 2), Article 101842. https://doi.org/10.1016/j.jormas.2024.101842

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & Prisma-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), Article 1. https://doi.org/10.1186/2046-4053-4-1

Mollaki, V. (2024). Death of a reviewer or death of peer review integrity? the challenges of using AI tools in peer reviewing and the need to go beyond publishing policies. *Research Ethics, 20*(2), 239–250. https://doi.org/10.1177/17470161231224552

Mondal, H., & Mondal, S. (2023). ChatGPT in academic writing: Maximizing its benefits and minimizing the risks. *Indian Journal of Ophthalmology, 71*(12), 3600–3606. https://doi.org/10.4103/IJO.IJO_718_23

Morocco-Clarke, A., Sodangi, F. A., & Momodu, F. (2023). The implications and effects of ChatGPT on academic scholarship and authorship: A death knell for original academic publications? *Information & Communications Technology Law, 33*(1), 21–41. https://doi.org/10.1080/13600834.2023.2239623

Nakavachara, V., Potipiti, T., & Chaiwat, T. (2024). Experimenting with generative AI: Does ChatGPT really increase everyone's productivity? *ArXiv*. https://doi.org/10.2139/ssrn.4746770

Nazzal, M. K., Morris, A. J., Parker, R. S., White, F. A., Natoli, R. M., Fehrenbacher, J. C., & Kacena, M. A. (2024). Using AI to write a review article examining the role of the nervous system on skeletal homeostasis and fracture healing. *Current Osteoporosis Reports, 22*, 217–221. https://doi.org/10.1007/s11914-023-00854-y

Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education, 49*(5), 847–864. https://doi.org/10.1080/03075079.2024.2323593

Olujimi, P. A., & Ade-Ibijola, A. (2023). NLP techniques for automating responses to customer queries: A systematic review. *Discover Artificial Intelligence, 3*, Article 20. https://doi.org/10.1007/s44163-023-00065-5

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., Hrobjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., …, & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Patil, N. S., Huang, R. S., van der Pol, C. B., & Larocque, N. (2024). Comparative performance of ChatGPT and bard in a text-based radiology knowledge assessment. *Canadian Association of Radiologists Journal, 75*(2), 344–350. https://doi.org/10.1177/08465371231193716

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator, 78*, 84–93. https://doi.org/10.1177/10776958221149577

Pérez-Núñez, A. (2023). Exploring the potential of generative AI (ChatGPT) for foreign language instruction: Applications and challenges. *Hispania, 106*, 355–362. https://doi.org/10.1353/hpn.2023.a906568

Perkins, M., & Roe, J. (2023). Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis. *F1000Research, 12*, Article 1398. https://doi.org/10.12688/f1000research.142411.2

Pesovski, I., Santos, R., Henriques, R., & Trajkovik, V. (2024). Generative AI for customizable learning experiences. *Sustainability, 16*, 3034. https://doi.org/10.3390/su16073034

Praveen, S. V., & Vajrobol, V. (2023). Understanding the perceptions of healthcare researchers regarding ChatGPT: A study based on bidirectional encoder representation from transformers (BERT) sentiment analysis and topic modeling. *Annals of Biomedical Engineering, 51*, 1654–1656. https://doi.org/10.1007/s10439-023-03222-0

Rane, N., Choudhary, S., & Rane, J. (2024). Intelligent manufacturing through generative artificial intelligence, such as ChatGPT or Bard. *SSRN*. https://doi.org/10.2139/ssrn.4681747

Rospigliosi, P. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments, 31*(1), 1–3. https://doi.org/10.1080/10494820.2023.2180191

Saad, A., Jenko, N., Ariyaratne, S., Birch, N., Iyengar, K. P., Davies, A. M., Vaishya, R., & Botchu, R. (2024). Exploring the potential of ChatGPT in the peer review process: An observational study. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 18*(2), Article 102946. https://doi.org/10.1016/j.dsx.2024.102946

Salimi, A., & Saheb, H. (2023). Large language models in ophthalmology scientific writing: Ethical considerations blurred lines or not at all? *American Journal of Ophthalmology, 254*, 177–181. https://doi.org/10.1016/j.ajo.2023.06.004

Salman, H., Al Mohsin, E., Al Rawi, A., & Shatnawi, S. (2022). Investigating HCI of the LMS blackboard ultra using WAMMI during COVID-19: Usability and design interactivity. In *Proceedings of the 2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies* (pp. 519–525). IEEE. https://doi.org/10.1109/3ICT56508.2022.9990650

Schmidt, L., Piazza, A., & Wiedenhöft, C. (2023). "Augmented brainstorming with AI"–Research approach for identifying design criteria for improved collaborative idea generation between humans and AI. In *Frontiers in artificial intelligence and applications* (pp. 410–412). IOS Press. https://doi.org/10.3233/FAIA230113

Semrl, N., Feigl, S., Taumberger, N., Bracic, T., Fluhr, H., Blockeel, C., & Kollmann, M. (2023). AI language models in human reproduction research: Exploring ChatGPT's potential to assist academic writing. *Human Reproduction, 38*(12), 2281–2288. https://doi.org/10.1093/humrep/dead207

Seth, I., Sinkjær Kenney, P., Bulloch, G., Hunter-Smith, D. J., Bo Thomsen, J., & Rozen, W. M. (2023). Artificial or augmented authorship? A conversation with a chatbot on base of thumb arthritis. *Plastic & Reconstructive Surgery-Global Open, 11*(5), Article e4999. https://doi.org/10.1097/GOX.0000000000004999

Victor, B. G., Sokol, R. L., Goldkind, L., & Perron, B. E. (2023). Recommendations for social work researchers and journal editors on the use of generative AI and large language models. *Journal of the Society for Social Work and Research, 14*, 563–577. https://doi.org/10.1086/726021

Watermeyer, R., Phipps, L., Lanclos, D., & Knight, C. (2023). Generative AI and the automating of academia. *Postdigital Science and Education, 6*, 446–466. https://doi.org/10.1007/s42438-023-00440-6

Wu, R. T., & Dang, R. R. (2023). ChatGPT in head and neck scientific writing: A precautionary anecdote. *American Journal of Otolaryngology, 44*(6), Article 103980. https://doi.org/10.1016/j.amjoto.2023.103980

Yan, L., Echeverria, V., Nieto, G. F., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2023). Human-AI collaboration in thematic analysis using ChatGPT: A user study and design recommendations. *ArXiv*. https://doi.org/10.1145/3613905.3650732

Zhang, P., & Kamel Boulos, M. N. (2023). Generative AI in medicine and healthcare: Promises, opportunities and challenges. *Future Internet, 15*(9), Article 286. https://doi.org/10.3390/fi15090286

Zheng, Z., Zhang, O., Borgs, C., Chayes, J. T., & Yaghi, O. M. (2023). ChatGPT chemistry assistant for text mining and the prediction of MOF synthesis. *Journal of the American Chemical Society, 145*, 18048–18062. https://doi.org/10.1021/jacs.3c05819

Zohouri, M., Sabzali, M., & Golmohammadi, A. (2024). Ethical considerations of ChatGPT-assisted article writing. *Synesis, 16*(1), 94–113.