SCIENTIFIC CULTURE

www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 2.1, (2026), pp. 12075-12090
Open Access. Online & Print

DOI: 10.5281/zenodo.20293644

GENERATIVE AI IN ACADEMIC RESEARCH: A STUDY OF IMPACT, CHALLENGES AND ETHICAL CONSIDERATIONS

Maidul Islam[1], Kuldeep Gujar[2*], Anshika Jangra[3], Shrabani Nanda[4] and Gi-Cheon Park[5]

[1]Department of International Business, Keimyung University, Daegu, South Korea
[2 5]Seoul Christian University, Seoul, South Korea
[3]The University of Suwon, Suwon, South Korea
[4, 5]Department of Business Intelligence, Ajou University, Suwon, South Korea

Received: 07/04/2026
Accepted: 08/05/2026

Corresponding Kuldeep Gujar
(dkekuldeep@gmail.com)

ABSTRACT

Recent advances in generative artificial intelligence have driven significant progress, and these gains in efficiency may substantially affect academic research. However, the exact role of generative artificial intelligence in research remains unclear and requires further study. Although artificial intelligence has a long research history, current developments such as Large Language Models present new roles that were not fully anticipated. This study addresses this gap by analysing the influence of generative artificial intelligence on research, examining how Large Language Models affect literature search, the research process, academic writing, peer review, and ethical issues. The primary objective is to evaluate the effectiveness of Large Language Models across various research activities. The paper also outlines the challenges and constraints these models introduce at different research stages. The contribution of this review lies in evaluating Large Language Models in research, examining their benefits and disadvantages, and exploring the ethical dimensions of their use. The study enhances the understanding of the roles, limitations, and contributions of Large Language Models to research. It also identifies several open research questions that warrant further scholarly investigation.

KEYWORDS: Generative AI; Large Language Models; Academic Research; Research Ethics; Generative AI in Research and Writing; ChatGPT; Peer Review; Literature Discovery

1. INTRODUCTION

Language models, such as ChatGPT, have significantly influenced research and academia. These technological advancements are reshaping scholarly practices, and researchers are exploring the impact of ChatGPT across various domains. This review paper examines this emerging phenomenon. The neural network architecture of ChatGPT and its large number of parameters are altering research dynamics, including academic writing, idea generation, research conduct, peer review, and editing. The rapid evolution of these models necessitates a comprehensive analysis of their implications within the academic landscape.

The study of the impact of ChatGPT spans multiple fields, including applications in training autonomous vehicles (Roumeliotis & Tselikas, 2023), aiding scientific idea generation (Rahman et al., 2023), gathering insights from healthcare researchers (Bélanger-Gravel et al., 2023), and potentially disrupting biomedical engineering (Ray et al., 2023). Large Language Models also excel in dialogue state tracking (Heck et al., 2023), business, and research (Gurjar et al., 2024). Given the breadth of these applications, ethical considerations necessitate thorough investigation (Huallpa, 2023). This paper contributes to the ongoing discourse by providing a focused analysis of the transformative effects of Large Language Models on core aspects of research and academia.

1.1. Large Language Models and Their Roles in Research

The following section provides an overview of Large Language Models and their evolving roles in reshaping research and academia. It also discusses the complexities and potential limitations of artificial intelligence systems in this context.

Large Language Models, particularly ChatGPT, are reshaping research and academia. In the legal field, these models may improve access to justice and increase efficiency in text retrieval and generation (Homoki & Ződi, 2024). In intelligence operations, they can assist officers in collecting, integrating, and producing reports (Snidaro, 2023). However, the complexities of artificial intelligence systems require caution and a nuanced understanding of their potential limitations (Tu et al., 2023). This study analyses existing literature to understand the potential of artificial intelligence to transform academic landscapes. It provides insights into efficient literature discovery, academic writing dynamics, and the peer review process. This investigation also considers other influential works to detail the impact of ChatGPT on specific aspects of research and academia, offering a holistic view of this evolving field.

1.2. Objectives and Contributions

The subsequent section outlines the core objectives and contributions of this study, detailing the key areas of investigation and the intended impact on the understanding of Large Language Models in research and academia.

The first objective is to evaluate the impact of Large Language Models on research and academia. This study investigates how artificial intelligence tools affect core research processes including data searching, referencing, and academic publishing. Examining the implications of artificial intelligence across domains provides a broad understanding of how these technologies have changed research practices, scholarly communication, and academic workflows.

The second objective is to analyse the influence of artificial intelligence on research conduct and idea generation. This research examines how artificial intelligence has shaped the way research is conducted, exploring the integration of artificial intelligence technologies into research methodologies and workflows. The analysis identifies changes that artificial intelligence brings, from data analysis and modelling to experimental design and hypothesis generation.

The third objective is to assess the role of Large Language Models in literature discovery. This study explores how effectively artificial intelligence tools help researchers find relevant papers. By analysing the capabilities and limitations of artificial intelligence-driven search engines, this study clarifies advancements in automating literature searches. It also provides insights into how researchers can use these tools to improve their discovery process.

The fourth objective is to investigate the evolution of academic writing in the era of artificial intelligence. This research examines the transformative effects of artificial intelligence on academic paper writing. Analysing artificial intelligence-assisted writing tools and language generation models highlights the changing landscape of academic writing, including improvements in productivity and language fluency, and the potential implications for authorship norms and writing conventions.

The fifth objective is to examine the impact of artificial intelligence on peer review. This study focuses on the influence of artificial intelligence on

traditional peer review systems. By analysing artificial intelligence-driven systems for peer review evaluations, it explores their potential benefits and challenges, including improved efficiency, objectivity, and quality assurance, as well as concerns about bias, transparency, and the role of human expertise.

The sixth objective is to evaluate the benefits, drawbacks, and ethical concerns of artificial intelligence in research. This study critically examines the benefits and drawbacks of using artificial intelligence tools in research, assessing the advantages of increased efficiency, accuracy, and scalability while also addressing concerns including data bias, interpretability challenges, and ethical considerations related to artificial intelligence-driven research practices.

2. BACKGROUND: LARGE LANGUAGE MODELS

Large Language Models are deep learning models that use a large number of parameters and are trained on vast amounts of text in an unsupervised manner (Alberts et al., 2023). Such models have developed rapidly since around 2018, with both the number of parameters and their capabilities increasing significantly. For example, GPT-4 has over one hundred trillion parameters and can process both text and visual information. ChatGPT, developed by OpenAI, is a notable Large Language Model that has gained considerable attention for its performance. ChatGPT is an artificial intelligence tool based on multilayer recurrent neural networks (Gokul, 2023), which are trained on extensive datasets. This training allows them to generate text that closely resembles human language. Unlike traditional language models, ChatGPT uses transformer-based architectures that enable parallel processing of large datasets (Gokul, 2023). Transformer-based models use neural networks that are known for fast and accurate data processing; they can generate substantial natural language output with minimal human intervention. Large Language Models can improve their performance using only unlabelled datasets (Tranfield et al., 2003).

From September 2022 to August 2023, the top ten artificial intelligence tools attracted over twenty-one billion visits (Sarkar, 2023). ChatGPT was the most popular, with 14.6 billion visits, representing over sixty per cent of the total visits. The artificial intelligence industry averaged three hundred million monthly visits, showing increasing user interaction. Generative artificial intelligence tools experienced a traffic increase of nearly 10.7 times from 2022 to 2023. This increase was driven by improved model capabilities, consumer interest, and wider applications. ChatGPT, Character AI, and Google Bard had significant net traffic increases, accumulating 1.8 billion, 463.4 million, and 68 million visits, respectively. In contrast, MidJourney and Quillbot experienced the most significant traffic declines during this period, as shown in Figure 1 and Figure 2. The United States accounted for 5.5 billion visits (22.62 per cent of the total). European countries collectively accounted for 3.9 billion visits. Artificial intelligence chatbot tools were the most popular, with 19.1 billion visits. Most users were male (69.5 per cent), compared to female users (30.5 per cent), and over sixty-three per cent of users accessed the platforms on mobile devices (Sarkar, 2023).

[Figure 1 – image placeholder; please insert original figure here]

Figure 1. Top 10 Large Language Models in use from September 2022 to August 2023 (Source: Sarkar, 2023).

To provide a more comprehensive understanding of Large Language Models, the following subsections detail their key characteristics and the architecture of an artificial intelligence agent system that incorporates them.

[Figure 2 – image placeholder; please insert original figure here]

Figure 2. The net traffic of the top 10 Large Language Models from September 2022 to August 2023.

2.1. Key Characteristics of Large Language Models

The following subsection outlines the key attributes that define and distinguish Large Language Models. These models exhibit several important characteristics:

- Large parameter size: Large Language Models are defined by their extensive number of parameters, which enable the model to learn complex patterns from data.
- Unsupervised training: these models are primarily trained on large amounts of unlabelled text data, and this unsupervised learning allows them to acquire broad language understanding.
- Contextual understanding: the models can understand the context of a given text input, which allows them to generate relevant and coherent responses.
- Text generation: they can produce human-

like text for various tasks, including writing, translation, and summarisation.

- Transformer architecture: many modern Large Language Models, including ChatGPT, use transformer networks. This architecture enables parallel processing, which significantly improves training efficiency.
- Emergent abilities: as Large Language Models scale in size, they demonstrate emergent abilities — capabilities that were not explicitly programmed, such as advanced reasoning and problem-solving.

2.2. Artificial Intelligence Agent System Architecture with Large Language Models

The following subsection describes how Large Language Models are integrated into artificial intelligence agent systems, including the essential components and their interactions. Large Language Models are often integrated into more complex systems such as artificial intelligence agents, which are designed to perform goal-directed tasks. Figure 3 illustrates a conceptual framework for such a system.

[Figure 3 — image placeholder; please insert original figure here]

Figure 3. Introduction to the primary features of autonomous multi-agent systems, utilising the capabilities of Large Language Models enriched by contextual resources such as tools and data (source: Händler, 2023).

This framework includes several key components. First, the user interaction layer manages how users interact with the system, involving the elicitation of prompts, decomposition of tasks, multi-agent collaboration, and alignment of system behaviour with user goals. Second, goal-driven task management allows Large Language Model-driven systems to manage complex tasks by breaking them down into smaller, more manageable subtasks. This process uses deep reasoning, or "slow thinking", facilitated by Large Language Model capabilities, with the system coordinating these subtasks to achieve the overall goal.

Third, intelligent agents form the core of these systems. These agents have specific competencies, defined roles, individual memories, and access to relevant resources, which can include data, tools, or other foundation models. Large Language Models enable these agents to reflect, plan, and efficiently process their assigned tasks, and agents can communicate with each other. Fourth, multi-agent collaboration allows Large Language Model-powered agents to work together in a network, exchanging messages, delegating responsibilities, seeking assistance, and evaluating the results of one another. This collaboration enables the achievement of more complex goals than individual agents could attain alone.

Fifth, context interaction is necessary because many tasks require agents to use contextual resources such as expert tools, data, or specialised foundation models. By leveraging these resources, agents can better understand and respond to their environment, enhancing their ability to handle diverse challenges. Finally, balancing autonomy and alignment is a key concern in Large Language Model-powered multi-agent systems. This balance is complex because it involves interactions between users, Large Language Model-powered agents, and governing mechanisms. System architecture must address this challenge to ensure efficiency and effectiveness, considering both general alignment principles and user-specific preferences. Optimising this balance is crucial for overall system performance.

2.3. Evaluation of Recent Large Language Models for Academic Writing

In this section, we provide a summary of some of the key papers in the field of generative artificial intelligence in academic writing. The selection of these papers is based on their relevance to our topic and their complementarity to our paper.

Recent research published by Aydin et al. (2025) has evaluated the performance of several Large Language Models in academic writing, providing insights into their potential and limitations. A study published in 2024 and 2025 compared the academic writing capabilities of DeepSeek v3, Qwen 2.5 Max, Qwen3 235B, ChatGPT, Gemini, Llama, Mistral, and Gemma (Aydin et al., 2025). This study addressed a gap in the literature regarding the extent to which these tools can be used to generate original content, focusing on content quality, readability, and effectiveness.

The evaluation involved generating texts from forty papers on Digital Twin and Healthcare. The Large Language Models produced content based on posed questions and paraphrased abstracts. The results of the study indicate several key trends.

- Plagiarism: paraphrased abstracts showed higher plagiarism rates, and question-based responses also exceeded acceptable plagiarism levels for academic work. However, Gemini 2.5 Pro and Qwen3 235B demonstrated more acceptable rates in question answering, while Llama 3.1 8B showed the lowest plagiarism rate in abstract paraphrasing. The

findings suggest that Large Language Models vary in their paraphrasing abilities and dependence on source material.

- Artificial intelligence detection: detection tools consistently identified nearly all generated texts as artificial intelligence-generated. While different tools (quillbot.com and stealthwriter.ai) showed variations in their measurements, the results were generally consistent. Some models were observed to leave fewer traces of artificial intelligence, producing more natural-sounding text.
- Word count: all chatbots produced a sufficient volume of content, with Qwen 2.5 Max and Qwen 3 235B generating the most words, particularly on knowledge-intensive topics. Mistral 7B and Deepseek-coder-v2 16B produced more concise outputs.
- Semantic similarity: paraphrased outputs from all models maintained a strong semantic relationship with the original texts, with similarity scores generally at ninety per cent or above. This indicates that the models effectively preserve semantic integrity during information transfer and paraphrasing.
- Readability: all models tended to produce texts that were complex and difficult to read. Sentences in both question-answer and paraphrased abstract texts were rated as very difficult to read by the Hemingway Editor. Gemini 2.5 Pro achieved the highest readability score on WebFX, but the scores were still low. The study suggests that Large Language Models tend to generate long, complex sentences with a wide vocabulary, including technical terms, which negatively impacts readability scores.

The study also compared its findings with previous research. The current study found relatively lower plagiarism rates compared to a prior study by Aydın and Karaarslan (2022), which evaluated the GPT-3 model of ChatGPT (Aydın & Karaarslan, 2022). The improvement in plagiarism rates may be attributed to advancements in Large Language Models in the intervening years. The authors conclude that while large and recent Large Language Models such as Qwen 3 235B, Qwen 2.5 Max, DeepSeek v3, and Gemini 2.5 Pro generally perform well in academic writing tasks, each model has different advantages and disadvantages depending on the specific usage scenario. Models such as Llama 3.1 8B and Llama 2 7B exhibit lower plagiarism rates and higher readability scores. Although some models can produce more natural-sounding text, artificial intelligence detection tools can still identify them. The findings offer guidance for optimising artificial intelligence-based text generation processes and selecting the most suitable model for a given context.

Another study by Dergaa et al. (2023) explored the potential of Natural Language Processing technologies, including ChatGPT, in academic writing and research publications (Dergaa et al., 2023). The research involved a literature review of scholarly articles, analysing the potential benefits, threats, and ethical considerations associated with these tools. The study identified the potential of ChatGPT and other Natural Language Processing technologies to enhance academic writing and research efficiency. However, it also raised concerns about the impact of these tools on the authenticity and credibility of academic work. The authors emphasised the necessity for comprehensive discussions regarding the potential uses, threats, and limitations of Natural Language Processing tools, advocating for the prioritisation of ethical and academic principles, with human intelligence and critical thinking remaining central to the research process. The paper concludes by calling for academics to address the challenges and opportunities presented by artificial intelligence, advocating for the education of research students on the ethical considerations and transparent use of these tools, and highlighting the importance of cross-checking artificial intelligence-generated information with reliable sources.

The study presented by Kim et al. (2025) explored students' perceptions and experiences with generative artificial intelligence-assisted academic writing, examining the expected roles, benefits, and challenges of generative artificial intelligence tools (Kim et al., 2025). Through interviews with Chinese university students, the research identified that students envision artificial intelligence as a multi-tasking writing assistant, virtual tutor, and digital peer, offering support across various writing processes. Students perceived benefits in the writing process, performance, and affective domain, but also noted challenges related to artificial intelligence, students themselves, and the tasks they undertake. The study emphasises the need for human-centred artificial intelligence design in education, advocating for a balanced symbiosis of artificial intelligence and human interaction, and proposes diversifying the roles of generative artificial intelligence to enhance instructional practices. It also highlights the importance of developing students' prompt engineering and higher-order thinking skills and calls for future research to explore these dynamics further, using larger samples, diverse learning tasks,

and longitudinal designs in actual classroom settings.

3. METHODOLOGY

As discussed earlier, the introduction of Large Language Models such as ChatGPT has significantly transformed research and academic writing. In this section, we review the methodology used for analysing research articles. Large Language Models have proven to be effective and powerful tools for researchers, with their capabilities in content creation and processing, and this has sparked a growing interest among researchers in investigating the impact of these models on research and academic writing. To conduct an in-depth review of the existing literature and bridge knowledge gaps, a systematic literature review method was employed. The systematic literature review method aims to address gaps in the current literature by answering unresolved questions and exploring future research areas. Additionally, bibliometric analysis was utilised to identify key trends, citation patterns, and thematic clusters within the literature. According to established standards, systematic reviews were conducted to address these gaps and assess the influence of Large Language Models on research (Jing et al., 2024; Antu et al., 2023; Paul & Criado, 2020). The systematic examination of the literature was carried out using the PRISMA model, which is recommended in the existing body of literature. This approach ensures that the review process is transparent and reproducible, and critically organises the research issues.

3.1. Selection of Databases

Various databases including Scopus, Web of Science, PubMed, IEEE Xplore, PsycINFO, ERIC, and Linguistics and Language Behavior Abstracts have been identified and examined by researchers. The Scopus and EBSCO databases were selected because they include and publish a wide range of research literature on the impact and influence of Large Language Models on research and academic writing. Scopus and EBSCO are among the largest abstract and citation databases, providing access to a wide repository of articles in journals, conference proceedings, and books, with broader coverage of peer-reviewed material. Researchers can locate relevant material by creating and customising their search queries, and various filters related to year, type of article, area, and search function can be applied to refine search results.

3.2. Selection of Articles

The search string includes keywords such as ("artificial intelligence" OR "AI" OR "AI tools" OR "chatbot" OR "bot" OR "chatbots" OR "machine intelligence" OR "automated intelligence" OR "Generative AI" OR "ChatGPT") and ("Research" OR "academic research" OR "academic" OR "academics") to search articles in the Scopus database and EBSCO. Similar words are used in the search string to expand the search area. Furthermore, the researchers examined both empirical and conceptual studies to refine the keyword search from recent review papers (Hulland, 2020; Bakiner, 2023). This comprehensive search yielded a broad range of papers relevant to the impact of artificial intelligence in research and academic writing. Figure 4 illustrates a PRISMA flow diagram that outlines the systematic literature search strategy employed, commencing with a defined keyword string across the Scopus and EBSCO databases, followed by the application of inclusion and exclusion criteria, ultimately yielding a final corpus of 129 eligible articles from which six key themes were synthesised for the review.

3.3. Inclusion and Exclusion Criteria

An inclusion and exclusion criterion was developed by considering numerous factors in order to select relevant and high-quality literature for the study. Only those studies meeting the standard were considered for the review process. With regard to inclusion, research articles published in peer-reviewed journals were eligible, including empirical studies and theoretical works that examine the influence of Large Language Models on research and academic writing. The studies considered examine the influence of Large Language Models on research and academic writing in the domains of education, information, decision sciences, information technology, computer science, and social sciences. Only articles published in the English language were considered for review. The search string had to appear either in the abstract, the title, or the author-listed keywords. Articles published by 31 December 2023 were considered for the review process.

With regard to exclusion, various document types such as editorials, letters, opinion pieces, and blog posts were excluded as they are not peer-reviewed. Articles that focused on core technical aspects such as model architecture and training techniques were excluded, as they do not focus on the applicability of Large Language Models in research and academic writing. Research articles not indexed in the Scopus and EBSCO databases were excluded from the analysis to focus on high-quality literature. Press articles were not included in the review. Retrieved

records were filtered according to predefined inclusion and exclusion criteria, which included relevance to the research topic, publication type, and language. After removing duplicate entries, the remaining articles were thoroughly assessed to determine their eligibility in the review process.

[Figure 4 – image placeholder; please insert original figure here]

Figure 4. PRISMA model with the search string.

3.4. Article Screening

The articles meeting the search criteria were screened and collected from the Scopus database and EBSCO. Figure 4 shows the PRISMA model displaying the search string used to reach the desired dataset. The keyword search string generated 125,802 potentially relevant papers. A robust data repository was developed by including the title and abstract of each article as it provides a thorough analysis of research articles. Duplicate articles were eliminated from the self-created article repository to obtain a clean dataset. In the further screening process, forty topics were identified by a set of keywords. Among those forty topics, fifteen themes were selected for further study. Manual screening and removal of publications from other topics were required to overcome algorithmic misclassification issues. Addressing algorithmic misclassification issues involved meticulously filtering out papers on academic writing, education, research, and other pertinent areas of academia. By conducting manual reviews, we ensured that the content aligned with the intended focus, enhancing the quality and relevance of the information available. The inclusion and exclusion criteria generated 221 articles which resulted in fifteen themes. Each member of the research team was assigned a topic to identify a suitable name that matches the theme. The grouped keywords helped to form clusters that generated appropriate themes. In the last stage, 129 research articles were considered eligible and generated six themes. The method used in selecting the article is dependable, visible, and error-free, and can produce recommendations based on the study (Rother, 2007).

Having established this robust and systematic approach for literature selection and categorisation, the subsequent section will now proceed to the analysis and synthesis of these identified themes and key findings, providing a comprehensive understanding of the current landscape of the impact of Large Language Models on research and academic writing.

4. ANALYSIS AND SYNTHESIS

This section synthesises current research on the impact of Large Language Models on academia. It analyses publication trends, key research areas, and prominent keywords to identify significant patterns and insights. This synthesis highlights the evolving landscape of Large Language Model research and its implications for the academic community. The analysis aims to provide a clear and structured understanding of how these models are shaping scholarly activities and future research directions.

4.1. Publication Trends

This subsection examines the year and journal of publication to illustrate the growth of Large Language Model research and its influence across academic disciplines. It details the expansion of artificial intelligence applications across various sectors and the corresponding increase in research output.

4.1.1. Artificial Intelligence Influence Across Sectors

Researchers extracted data on the year and journal of publication, focusing on articles related to Large Language Models in research. Figure 5 illustrates the influence of artificial intelligence across various sectors. Artificial intelligence has significantly penetrated numerous sectors, transforming both research and industry practices. In healthcare, approximately sixty per cent of organisations leverage artificial intelligence to enhance diagnostics, patient care, and drug discovery (Singla et al., 2024). The finance industry also exhibits a high artificial intelligence utilisation rate of sixty per cent, particularly in fraud detection, algorithmic trading, and customer support, thereby enhancing operational efficiency and risk management. Artificial intelligence adoption is near forty-seven per cent in retail and e-commerce, where it supports personalised shopping experiences and optimises inventory management (Singla et al., 2024).

In manufacturing and energy, artificial intelligence usage stands at thirty-eight per cent and thirty per cent, respectively. Artificial intelligence aids in predictive maintenance, quality control, and energy grid optimisation (Thormundsson, 2024). These applications increase resource efficiency and minimise downtime. Agriculture and environmental science employ artificial intelligence at rates of around twenty per cent and thirty per cent, respectively, for applications such as crop yield prediction, pest control, and environmental modelling. Other sectors, including

telecommunications (thirty-five per cent), security and defence (forty-five per cent), and human resources (thirty-five per cent), utilise artificial intelligence for network optimisation, cybersecurity measures, and employee management. Artificial intelligence is also transforming legal services (ten per cent), real estate (fifteen per cent), and media (twenty per cent) by automating tasks, providing data-driven insights, and enhancing user engagement (Singla et al., 2024; Thormundsson, 2024).

[Figure 5 – image placeholder; please insert original figure here]

Figure 5. Research areas of artificial intelligence influence.

4.1.2. Growth of Artificial Intelligence Publications

Figure 6 demonstrates a substantial increase in annual artificial intelligence-related journal articles from 2013 to 2021. The number of publications rose from 298,381 in 2013 to 861,450 in 2021. This growth reflects the escalating interest in artificial intelligence research, driven by advances in machine learning, deep learning, and related technologies. The expansion of artificial intelligence applications across sectors such as healthcare, finance, and manufacturing has also fuelled this trend. The period from 2018 to 2021, in particular, shows a marked increase in artificial intelligence research investment.

[Figure 6 – image placeholder; please insert original figure here]

Figure 6. Yearly number of journal articles published on the topic of artificial intelligence.

This surge was likely influenced by significant technological breakthroughs and the expansion of the global artificial intelligence research community. Increased funding from both government and private sectors, coupled with the growing need for advanced artificial intelligence solutions, has also contributed to this expansion (Giattino et al., 2023; Maslej et al., 2023).

4.1.3. Regional Distribution of Artificial Intelligence Publications

Figure 7 illustrates the global distribution of artificial intelligence journal publications from 2010 to 2021. Artificial intelligence research publications were primarily concentrated in East Asia and the Pacific (47.14 per cent), largely due to the significant research output of China. North America followed with 11.61 per cent, driven by contributions from academic and technology sectors in the United States. Europe and Central Asia accounted for 7.20 per cent, supported by EU-funded artificial intelligence initiatives. South Asia, led mainly by India, contributed 6.75 per cent, reflecting its increasing investment in artificial intelligence. The Middle East and North Africa (4.64 per cent) and Sub-Saharan Africa (0.77 per cent) exhibited lower publication outputs, indicating regional disparities in artificial intelligence research intensity (Maslej et al., 2023). Governments worldwide are encouraging researchers to contribute to artificial intelligence policy development. India ranks fourth in artificial intelligence research output. The Indian government has promoted artificial intelligence through initiatives such as the National AI Portal and the National AI Strategy. Agencies such as the Department of Science and Technology and the Ministry of Electronics and Information Technology actively support research collaboration between industry and academia.

[Figure 7 – image placeholder; please insert original figure here]

Figure 7. Regional distribution of artificial intelligence publications, 2010–2021.

4.2. Keyword Analysis

This subsection explores the prominent research keywords used in the analysed publications to provide insight into the key themes and focus areas within Large Language Model research. Keyword analysis, conducted using VOSviewer, identified key terms such as "artificial intelligence", "human", "ethics", "research", and "ChatGPT". Figure 8 illustrates the relationships between these keywords. Thicker lines indicate a greater volume of research on a particular theme, while thinner lines suggest relatively less explored areas. The clusters highlight that artificial intelligence ethics, ethical issues, and privacy are areas requiring further research. The clusters also reveal that artificial intelligence innovation in research and development is a prominent topic. The strength of the connections between keywords indicates the degree of emphasis within the research. Areas such as the use of artificial intelligence in higher education and artificial intelligence ethics are significant and warrant more in-depth investigation.

[Figure 8 – image placeholder; please insert original figure here]

Figure 8. Keyword analysis.

5. THEMES AROUND LARGE LANGUAGE MODELS IN RESEARCH

Previous sections described the influence of artificial intelligence in research and academia, including the use of ChatGPT. This section analyses and synthesises these contributions, focusing on four key dimensions: the impact of artificial intelligence tools on research and academics, the discovery of relevant papers, how Large Language Models are changing research conduct and idea generation, and the transformative impact of artificial intelligence on academic writing and the peer review process.

The recent rise of artificial intelligence and ChatGPT has generated significant interest among researchers and educators in integrating these tools into research. Artificial intelligence-based tools assist researchers in conducting, analysing, assessing, and writing research. Machine learning and Natural Language Processing have been incorporated into many applications to facilitate research, writing, critical analysis, literature reviews, and referencing (Kitchenham et al., 2009). Although researchers are eager to adopt artificial intelligence tools, their application, strengths, and weaknesses require further investigation.

A central concern in using Large Language Models in education and research is balancing efficiency and authenticity in machine-generated content. Artificial intelligence can quickly produce substantial amounts of text and other content, which is useful for academics and researchers. However, this raises the risk of sacrificing the distinctiveness, artistry, and creativity of human writing. Research suggests that adopting responsible and ethical standards in academia and research is crucial for maximising the benefits of artificial intelligence and mitigating potential drawbacks (Snyder, 2019). While ChatGPT is perceived by some as a threat that could normalise plagiarism and endanger scientific research, others argue that it can enhance the quality and legitimacy of work by drawing on diverse data sources (Snyder, 2019). Addressing the challenges posed by artificial intelligence tools in research is essential (Webster & Watson, 2002). Potential responses include developing anti-ChatGPT software, encouraging more creative and innovative research, and reconsidering the traditional "publish or perish" model. This section analyses four major domains where generative artificial intelligence impacts academic research. Ethical considerations, although pertinent to each domain, are discussed in detail in Section 6.

5.1. The Influence of Large Language Models on Research Conduct and Idea Generation

Large Language Models such as ChatGPT have significantly influenced research, prompting varied perspectives on their role and potential to enhance research. They can improve information aggregation and interdisciplinary collaboration, thereby increasing scholarly impact (Dalalah & Dalalah, 2023). However, the integration of artificial intelligence into research processes, especially for administrative tasks, warrants careful consideration. Overreliance on utilitarian applications may compromise the richness of academic culture and creative scholarly pursuits. Artificial intelligence-driven methodologies can optimise research, enhancing efficiency and outcomes. The integration of artificial intelligence tools in research libraries streamlines critical operations such as cataloguing, information retrieval, and knowledge organisation, benefiting the scholarly community (Titko et al., 2023). Furthermore, artificial intelligence automates research funding applications and proposal preparation (Piwowar et al., 2018). By identifying suitable funding sources, automating administrative tasks, and facilitating proposal development, artificial intelligence can facilitate the funding process and foster innovative scholarship.

In summary, artificial intelligence tools are becoming integral to the evolving research landscape, offering transformative possibilities. While artificial intelligence can expedite and enhance research, it is crucial to proceed cautiously, addressing ethical concerns, ensuring proper oversight, and upholding scholarly integrity. Institutional Review Boards play a vital role in protecting the rights and welfare of human research subjects (Chadwick et al., 2000), especially given increasing research pressures.

5.2. Large Language Models and the Discovery of Relevant Papers

The effectiveness of artificial intelligence tools in facilitating literature searches has been widely acknowledged in scholarly discourse. For example, Research Rabbit, a literature mapping tool, helps researchers navigate the complexities of scholarly literature (Denyer & Tranfield, 2009). This tool uses artificial intelligence algorithms to guide researchers through scholarly articles, enabling them to quickly identify relevant papers and extract key insights. Machine learning algorithms also show great potential in literature reviews (Tu et al., 2023). An empirical study demonstrated that artificial

intelligence-driven methodologies achieved a ninety per cent recall rate with only six per cent of the effort required by traditional manual methods. This highlights the efficiency and time-saving benefits of artificial intelligence tools in literature discovery.

Natural Language Processing techniques also enhance information retrieval and monitoring (Prokhorov & Safronov, 2019). Such methodologies can improve the accuracy and efficiency of information retrieval, providing researchers with faster access to relevant papers. Another study compared systematic and semi-systematic literature reviews, introducing a novel tool that integrates various artificial intelligence applications to optimise the research process by increasing speed, quality, and cost-efficiency (Müller et al., 2022). Artificial intelligence techniques such as machine learning, data mining, and text analytics streamline the literature search process, enabling researchers to more efficiently identify pertinent papers. Research confirms that artificial intelligence tools can significantly enhance the efficiency and effectiveness of literature searches. The integration of artificial intelligence algorithms, machine learning, and Natural Language Processing has effectively retrieved relevant papers while reducing the labour-intensive nature of traditional approaches. Artificial intelligence tools offer an auspicious alternative to conventional paper discovery methods, providing researchers with innovative, comprehensive, and expeditious means to conduct their work.

Information retrieval and academic literature discovery systems dynamically respond to user actions by presenting articles that align with the information needs of the user, similar to recommender systems in computer science (Kurtz & Henneken, 2014). For instance, researchers have aimed to identify and summarise exemplary research, such as big data and smart health strategies papers published in 2013 (Koutkias & Thiessard, 2014). They employed a systematic search method, examining journal papers in major bibliographic databases. Additionally, a comprehensive review of common datasets used in existing literature has contributed to the field (Obidallah et al., 2020). Challenges remain in identifying relevant literature, particularly when similar research domains use different terminology. In 2021, a system called VITALITY was introduced to address these challenges and improve current practices (Narechania et al., 2021). This retrieval-based paradigm demonstrates the potential of Large Language Models in information retrieval and academic literature discovery. Large Language Models have also garnered research attention for their strong performance in text-based tasks. A 2023 study categorised and reviewed existing Large Language Model research from the perspective of seven major software engineering tasks (Zheng et al., 2023). This analysis helps researchers navigate research trends and address the challenges of applying Large Language Models in various contexts. These efforts, along with other influential contributions, have enriched information retrieval and literature discovery, advancing the field and providing researchers with innovative tools and methodologies (Ellaway & Tolsgaard, 2023).

5.3. The Transformative Impact of Large Language Models on Academic Writing and Scholarly Publishing

Large Language Models have significantly transformed academic writing, fundamentally altering scholarly literature and the processes of manuscript submission and revision. This transformation is evident in several key developments. In 2019, the role of artificial intelligence in efficiently identifying suitable peer reviewers, streamlining the peer review process, and enhancing academic publishing was emphasised (Upshall, 2019). More recently, the introduction of the Manubot publishing ecosystem in 2023 demonstrated how artificial intelligence can collaboratively suggest revisions, improving the quality and narrative coherence of scholarly writing (Pividori & Greene, 2023).

A 2021 survey of artificial intelligence tools in publishing predicted a radical transformation of the industry (Razack et al., 2021). These advancements enhance various publishing functions and responsibilities. Another research group highlighted scalability, noting the role of artificial intelligence in assessing the quality of academic submissions, thereby improving overall quality and knowledge dissemination (Zhu et al., 2021). These artificial intelligence-driven advancements are potentially transforming academic writing and publishing, ushering in a technology-driven era characterised by improved content, fewer errors, and more seamless knowledge dissemination. The use of Large Language Models in education, particularly in digital and essay writing and other academic tasks, has also been explored. The potential of these models to generate original and coherent text raises concerns about academic integrity (Perkins, 2023). Decisions regarding the use of Large Language Models by students should be guided by updated academic integrity policies.

The role of ChatGPT in academic writing, including material organisation, draft creation, and proofreading, is significant (Altmäe et al., 2023). However, the potential threat of artificial intelligence to academic integrity in educational settings has raised concerns. One study evaluated the performance of ChatGPT in complex scientific writing tasks (Williams & Fadda, 2023). The evolving landscape has also prompted discussions on authorship and the appropriate use of artificial intelligence in academic papers. A framework called "PaperCard" was proposed to transparently declare the role of artificial intelligence in the writing process (Cho et al., 2023). Additionally, a human–artificial intelligence collaboration framework for artificial intelligence in writing was introduced, and the potential benefits and ethical considerations of artificial intelligence tools in academic writing were explored (Dergaa et al., 2023; Lin, 2023). The need to distinguish between human and artificial intelligence-generated content in academic writing, especially in higher education, is urgent (Desaire et al., 2023). The feasibility of using ChatGPT in academic writing within human reproduction research has also been examined (Semrl et al., 2023). Recent studies highlight the far-reaching impact of Large Language Models on academic writing, sparking discussions on ethics, authorship, and quality (Semrl et al., 2023).

5.4. Impact of Large Language Models on the Peer Review Process

The peer review process is essential to scholarly publishing, ensuring the integrity and quality of research. This section examines the transformative influence of artificial intelligence on this process, focusing on how artificial intelligence-driven systems can enhance efficiency and decision-making.

The traditional peer review process involves submitting a manuscript to a journal, where expert reviewers evaluate its scientific rigour and validity. Their insights are invaluable for refining scholarly discourse. However, this human-driven process can lead to significant publication delays. An artificial intelligence-assisted decision support system was developed to provide timely guidance to editors, reviewers, and authors, expediting manuscript evaluations through intelligent recommendations (Ghosal, 2019).

Researchers advocate for semi-automated peer review systems that leverage the analytical capabilities of artificial intelligence (Checco et al., 2021). These systems can accurately identify potential research weaknesses and low-quality or contentious studies. They can also match reviewers with appropriate expertise, ensuring thorough evaluations. Artificial intelligence-driven systems have the potential to enhance the precision and dependability of peer review. The impact of writing initiatives on scholarly output has been assessed, highlighting the benefits of collaborative learning and peer review (Kulage & Larson, 2016). Further research is needed to fully understand the long-term impact of peer review on academic essays (de Brusa & Harutyunyan, 2019).

6. LIMITATIONS AND ETHICAL CONCERNS OF USING LARGE LANGUAGE MODELS IN RESEARCH

This section explains the limitations and ethical concerns related to the use of Large Language Models in research and academia. It is divided into two subsections.

6.1. Limitations of Large Language Models

This section examines the limitations of Large Language Models, focusing on ChatGPT, in healthcare education, research, and practice. While Large Language Models such as ChatGPT offer advantages, they also present challenges that may hinder their integration into educational and research settings. This analysis draws on recent research that investigates the utility of these models and emphasises the need to recognise and address their inherent limitations.

Large Language Models such as ChatGPT may cite nonexistent scientific sources, and these inaccuracies are difficult to correct. ChatGPT can be used for both constructive and detrimental purposes, posing a risk of misuse (Akiba & Fraboni, 2023). Responsible use of ChatGPT is essential, given its susceptibility to exploitation.

Using ChatGPT for academic paper composition involves inherent risks. It can quickly produce a draft that might take weeks or months for an academic to complete. While this time-saving feature is appealing, caution is necessary. ChatGPT uses a vast database from various websites, and the origin of its information is unclear. Requested references may include fabricated DOIs or unrelated articles. Therefore, authors must manually fact-check the generated information to ensure accuracy. This step is crucial for maintaining academic integrity and the reliability of research.

6.2. Ethical Concerns of Large Language Models

This section examines the ethical dimensions surrounding Large Language Models in research,

addressing concerns about their potential implications. As Large Language Models such as ChatGPT are increasingly used across domains, their ethical considerations require scrutiny. The discussion is based on key works that explore the ethical challenges and potential risks associated with these models, highlighting the need for their responsible development, deployment, and use.

The AI SOCRATES system, based on Socratic philosophy, was introduced to address ethical dilemmas through deliberative dialogue, moving beyond simple answers (Bang et al., 2022). A systematic literature review identified the current state of Large Language Model research in automating educational tasks and offered recommendations for future work (Bang et al., 2022). An overview of the security risks of ChatGPT emphasised the ethical and security implications that require ongoing investigation (Derner & Batistič, 2023). Proposed domains for journals consider the ethical use of Large Language Models in research (Salimi & Saheb, 2023). The development of a citation mechanism was advocated, despite potential pitfalls, and research problems were outlined to promote responsible use of these models (Huang & Chang, 2023). Challenges, including ethical considerations, in the interaction of information retrieval models, Large Language Models, and human evaluators were highlighted (Ai et al., 2023). The ethical and technical challenges posed by Large Language Models were addressed, urging researchers to tackle issues such as privacy, data security, and potential biases (Romano et al., 2023). Various studies continue to contribute to the ongoing discourse on ethical considerations (Ellaway & Tolsgaard, 2023; Strasser, 2023). The implementation of language models such as ChatGPT raises ethical and regulatory issues, including data privacy, data security, and potential biases (Schwartz & Kalichman, 2007). Ethical issues in big data were addressed through a post-humanist framework, focusing on power relationships and potential harm to research participants (Mauthner, 2019). The potential of ChatGPT in aiding clinical text mining was explored, with an emphasis on the uncertain effectiveness of these models in healthcare settings (Tang et al., 2023). These studies collectively underscore the complex ethical landscape surrounding Large Language Models and the importance of ongoing scholarly debate.

7. DISCUSSION AND FUTURE WORK

It is essential to verify the originality of artificial intelligence-generated content and detect any plagiarism. Research should prioritise assessing the reliability of artificial intelligence-generated materials, focusing on potential errors, biases, and misinformation. Establishing clear protocols for integrating artificial intelligence in academic publishing is crucial, including procedures for transparently acknowledging the use of artificial intelligence, recognising its limitations, and validating data sources (Dupps, 2023). Further research should thoroughly evaluate the accuracy and reliability of Large Language Model tools. Additionally, exploring how these tools can enhance student support beyond academic writing is promising. This exploration should also consider the potential impact on pedagogy and learning outcomes.

The future of Large Language Models is complex and influenced by many factors, making it difficult to categorise the current market stage. This section presents several potential explanations for the uncertainties surrounding the future of these models.

Regulatory changes are a significant variable. Governments may introduce strict regulations that impact artificial intelligence companies, potentially hindering innovation and user engagement. Stricter data privacy laws, for example, might limit data collection, affecting artificial intelligence services that rely on user data. The evolving legal landscape necessitates careful consideration of these factors.

Economic factors may also affect the artificial intelligence landscape. Despite a substantial increase in interest (10.7 times compared to the previous year), the initial surge of enthusiasm and exploration may have subsided. This could indicate industry maturation, with users transitioning from exploration to more focused and sustained usage. The long-term economic sustainability of Large Language Model development and deployment remains a key area for analysis.

Shifts in consumer preferences can significantly impact Large Language Models. Rapid changes in trends and public opinion, especially regarding ethical considerations, could alter user trust or preferences for non-artificial intelligence solutions. These shifts may influence user engagement and the use of artificial intelligence tools. Understanding these evolving preferences is crucial for the responsible development and adoption of Large Language Models.

The adoption of mobile applications by artificial intelligence companies adds complexity. As artificial intelligence services develop mobile applications, users may prefer accessing them on their phones rather than through web browsers. While this shift may indicate positive engagement, it could also lead

to a decline in web traffic. This highlights the importance of multi-platform accessibility and user experience design.

8. CONCLUSIONS

In conclusion, this review has thoroughly explored the impact of Large Language Models on research and academia. The analysis highlights the transformative potential of artificial intelligence in reshaping academic research, including literature searches, research methodologies, academic writing, and peer review. Large Language Model tools can optimise literature searches, provide targeted information, and enhance search efficiency. They can also improve writing quality, identify errors, and transform scholarly publishing. However, ethical considerations and concerns about potential biases require careful attention and resolution. Although artificial intelligence-driven systems can streamline peer review, human judgement remains essential. Promoting the responsible integration of artificial intelligence in research practices is crucial. This will ensure that its beneficial capabilities are used to uphold integrity and equity in research. Future research should continue to examine the evolving role of Large Language Models in shaping the future of academia.

Author Contributions: Conceptualization, K.G., A.J., S.N., M.I. and G.-C.P.; methodology, K.G.; validation, M.I. and G.-C.P.; formal analysis, K.G., A.J., S.N. and M.I.; investigation, S.N.; resources, M.I.; data curation, K.G., A.J. and S.N.; writing—original draft preparation, K.G., A.J. and S.N.; writing—review and editing, K.G., A.J. and S.N.; visualisation, M.I. and K.G.; supervision, G.-C.P.; funding acquisition, M.I. All authors have read and agreed to the published version of the manuscript.

Acknowledgements: The authors thank the anonymous reviewers for their constructive comments. The Article Processing Charge was partially funded by Keimyung University, Daegu, South Korea. The authors declare no conflicts of interest.

REFERENCES

Ai, Q., Bai, T., Cao, Z., Chang, Y., Chen, J., Chen, Z., et al. (2023) Information retrieval meets large language models: A strategic report from the Chinese IR community. AI Open, Vol. 4, 80–90.

Akiba, D. and Fraboni, M. C. (2023) AI-supported academic advising: Exploring ChatGPT's current state and future potential toward student empowerment. Education Sciences, Vol. 13, No. 9, 885.

Alberts, I. L., Mercolli, L., Pyka, T., Prenosil, G., Shi, K., Rominger, A., et al. (2023) Large language models (LLM) and ChatGPT: What will the impact on nuclear medicine be? European Journal of Nuclear Medicine and Molecular Imaging, Vol. 50, No. 6, 1549–1552.

Altmäe, S., Sola-Leyva, A. and Salumets, A. (2023) Artificial intelligence in scientific writing: A friend or a foe? Reproductive BioMedicine Online, Vol. 47, No. 1, 3–9.

Antu, S. A., Chen, H. and Richards, C. K. (2023) Using LLM (large language model) to improve efficiency in literature review for undergraduate research. Proceedings of the LLM@AIED Workshop, 8–16.

Aydin, O., Karaarslan, E., Erenay, F. S. and Bacanin, N. (2025) Generative AI in academic writing: A comparison of DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma. arXiv preprint, arXiv:2503.04765.

Aydın, Ö. and Karaarslan, E. (2022) OpenAI ChatGPT generated literature review: Digital twin in healthcare. In Aydın, Ö. (Ed.) Emerging Computer Technologies, Vol. 2, 22–31.

Bakiner, O. (2023) What do academics say about artificial intelligence ethics? An overview of the scholarship. AI and Ethics, Vol. 3, No. 2, 513–525.

Bang, Y., Lee, N., Yu, T., Khalatbari, L., Xu, Y., Su, D., et al. (2022) AiSocrates: Towards answering ethical quandary questions. arXiv preprint, arXiv:2205.05989.

Bélanger-Gravel, A., Janezic, I., Desroches, S., Paquette, M.-C., Therrien, F., Barnett, T., et al. (2023) Examining public health practitioners' perceptions and use of behavioural sciences to design health promotion interventions. BMC Health Services Research, Vol. 23, No. 1, 1–10.

Chadwick, P., Lees, S. and Birchwood, M. (2000) The revised Beliefs about Voices Questionnaire (BAVQ–R). The British Journal of Psychiatry, Vol. 177, No. 3, 229–232.

Checco, A., Bracciale, L., Loreti, P., Pinfield, S. and Bianchi, G. (2021) AI-assisted peer review. Humanities and Social Sciences Communications, Vol. 8, No. 1, 1–11.

Cho, W. I., Cho, E. and Cho, K. (2023) PaperCard for reporting machine assistance in academic writing. arXiv preprint, arXiv:2310.04824.

Dalalah, D. and Dalalah, O. M. (2023) The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. The International Journal of Management Education, Vol. 21, No. 2, 100822.

de Brusa, M. F. P. and Harutyunyan, L. (2019) Peer review: A tool to enhance the quality of academic written productions. English Language Teaching, Vol. 12, No. 5, 30–39.

Denyer, D. and Tranfield, D. (2009) Producing a systematic review. In Buchanan, D. A. and Bryman, A. (Eds.) The Sage Handbook of Organizational Research Methods. London, Sage Publications.

Dergaa, I., Chamari, K., Zmijewski, P. and Saad, H. B. (2023) From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. Biology of Sport, Vol. 40, No. 2, 615–622.

Derner, E. and Batistič, K. (2023) Beyond the safeguards: Exploring the security risks of ChatGPT. arXiv preprint, arXiv:2305.08005.

Desaire, H., Chua, A. E., Isom, M., Jarosova, R. and Hua, D. (2023) ChatGPT or academic scientist? Distinguishing authorship with over 99 % accuracy using off-the-shelf machine learning tools. arXiv preprint, arXiv:2303.16352.

Dupps, W. J., Jr. (2023) Artificial intelligence and academic publishing. Journal of Cataract & Refractive Surgery, Vol. 49, No. 7, 655–656.

Ellaway, R. H. and Tolsgaard, M. (2023) Artificial scholarship: LLMs in health professions education research. Advances in Health Sciences Education, Vol. 28, 1–6.

Ghosal, T. (2019) Exploring the implications of artificial intelligence in various aspects of scholarly peer review. Bulletin of the IEEE Technical Committee on Digital Libraries, Vol. 15, No. 1.

Giattino, C., Mathieu, E., Samborska, V. and Roser, M. (2023) Annual scholarly publications on artificial intelligence. Our World in Data. Available at https://ourworldindata.org/grapher/annual-scholarly-publications-on-artificial-intelligence.

Gokul, A. (2023) LLMs and AI: Understanding its reach and impact. Working Paper.

Gurjar, K., Jangra, A., Baber, H., Islam, M. and Sheikh, S. A. (2024) An analytical review on the impact of artificial intelligence on the business industry: Applications, trends, and challenges. IEEE Engineering Management Review.

Händler, T. (2023) A taxonomy for autonomous LLM-powered multi-agent architectures. Proceedings of the 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. doi:10.5220/0012239100003598.

Heck, M., Lubis, N., Ruppik, B., Vukovic, R., Feng, S., Geishauser, C., et al. (2023) ChatGPT for zero-shot dialogue state tracking: A solution or an opportunity? arXiv preprint, arXiv:2306.01386.

Homoki, P. and Ződi, Z. (2024) Large language models and their possible uses in law. Hungarian Journal of Legal Studies.

Huallpa, J. J. (2023) Exploring the ethical considerations of using Chat GPT in university education. Periodicals of Engineering and Natural Sciences, Vol. 11, No. 4, 105–115.

Huang, J. and Chang, K. C.-C. (2023) Citation: A key to building responsible and accountable large language models. arXiv preprint, arXiv:2307.02185.

Hulland, J. (2020) Conceptual review papers: Revisiting existing research to develop and refine theory. AMS Review, Vol. 10, No. 1, 27–35.

Jing, Y., Wang, C., Chen, Y., Wang, H., Yu, T. and Shadiev, R. (2024) Bibliometric mapping techniques in educational technology research: A systematic literature review. Education and Information Technologies, Vol. 29, No. 8, 9283–9311.

Kim, J., Yu, S., Detrick, R. and Li, N. (2025) Exploring students' perspectives on generative AI-assisted academic writing. Education and Information Technologies, Vol. 30, No. 1, 1265–1300.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J. and Linkman, S. (2009) Systematic literature reviews in software engineering — A systematic literature review. Information and Software Technology, Vol. 51, No. 1, 7–15.

Koutkias, V. and Thiessard, F. (2014) Big data — Smart health strategies. Yearbook of Medical Informatics, Vol. 23, No. 1, 48–51.

Kulage, K. M. and Larson, E. L. (2016) Implementation and outcomes of a faculty-based, peer review manuscript writing workshop. Journal of Professional Nursing, Vol. 32, No. 4, 262–270.

Kurtz, M. J. and Henneken, E. A. (2014) Finding and recommending scholarly articles. In Cronin, B. and

Sugimoto, C. R. (Eds.) Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact. Cambridge, MA, MIT Press, 243–259.

Lin, Z. (2023) Supercharging academic writing with generative AI: Framework, techniques, and caveats. arXiv preprint, arXiv:2310.17143.

Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Parli, V., Shoham, Y., Wald, R., Clark, J. and Perrault, R. (2023) The AI Index 2023 Annual Report. Stanford, CA, AI Index Steering Committee, Stanford University.

Mauthner, N. S. (2019) Toward a posthumanist ethics of qualitative research in a big data era. American Behavioral Scientist, Vol. 63, No. 6, 669–698.

Müller, H., Pachnanda, S., Pahl, F. and Rosenqvist, C. (2022) The application of artificial intelligence on different types of literature reviews — A comparative study. Proceedings of the 2022 International Conference on Applied Artificial Intelligence (ICAPAI), 1–7.

Narechania, A., Karduni, A., Wesslen, R. and Wall, E. (2021) VITALITY: Promoting serendipitous discovery of academic literature with transformers and visual analytics. IEEE Transactions on Visualization and Computer Graphics, Vol. 28, No. 1, 486–496.

Obidallah, W. J., Raahemi, B. and Ruhi, U. (2020) Clustering and association rules for web service discovery and recommendation: A systematic literature review. SN Computer Science, Vol. 1, 1–33.

Paul, J. and Criado, A. R. (2020) The art of writing literature review: What do we know and what do we need to know? International Business Review, Vol. 29, No. 4, 101717.

Perkins, M. (2023) Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching & Learning Practice, Vol. 20, No. 2, 7.

Pividori, M. and Greene, C. S. (2023) A publishing infrastructure for AI-assisted academic authoring. bioRxiv preprint, doi:10.1101/2023.01.21.525030.

Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B., et al. (2018) The state of OA: A large-scale analysis of the prevalence and impact of open access articles. PeerJ, Vol. 6, e4375.

Prokhorov, S. and Safronov, V. (2019) AI for AI: What NLP techniques help researchers find the right articles on NLP. Proceedings of the 2019 International Conference on Artificial Intelligence: Applications and Innovations (IC-AIAI), 76–765.

Rahman, M. M., Terano, H. J., Rahman, M. N., Salamzadeh, A. and Rahaman, M. S. (2023) ChatGPT and academic research: A review and recommendations based on practical examples. Journal of Education, Management and Development Studies, Vol. 3, No. 1, 1–12.

Ray, P., Chakraborty, R., Banik, O., Banoth, E. and Kumar, P. (2023) Surface engineering of a bioartificial membrane for its application in bioengineering devices. ACS Omega, Vol. 8, No. 4, 3606–3629.

Razack, H. I. A., Mathew, S. T., Saad, F. F. A. and Alqahtani, S. A. (2021) Artificial intelligence-assisted tools for redefining the communication landscape of the scholarly world. Science Editing, Vol. 8, No. 2, 134–144.

Romano, M. F., Shih, L. C., Paschalidis, I. C., Au, R. and Kolachalama, V. B. (2023) Large language models in neurology research and future practice. Neurology, Vol. 101, No. 23, 1058–1067.

Rother, E. T. (2007) Systematic literature review × narrative review. Acta Paulista de Enfermagem, Vol. 20, v–vi.

Roumeliotis, K. I. and Tselikas, N. D. (2023) ChatGPT and Open-AI models: A preliminary review. Future Internet, Vol. 15, No. 6, 192.

Salimi, A. and Saheb, H. (2023) Large language models in ophthalmology scientific writing: Ethical considerations blurred lines or not at all? American Journal of Ophthalmology, Vol. 254, 177–181.

Sarkar, S. (2023) AI industry analysis: 50 most visited AI tools and their 24B+ traffic behaviour. Writerbuddy, 4 December.

Schwartz, P. H. and Kalichman, M. W. (2007) Ethical concerns for stem cell research. In Human Stem Cell Manual. Amsterdam, Elsevier, 426–436.

Semrl, N., Feigl, S., Taumberger, N., Bracic, T., Fluhr, H., Blockeel, C., et al. (2023) AI language models in human reproduction research: Exploring ChatGPT's potential to assist academic writing. Human Reproduction, Vol. 38, No. 12, 2281–2289.

Singla, A., Yee, L., Chui, M. and Hall, B. (2024) The state of AI in early 2024: Gen AI adoption spikes and starts to generate value. McKinsey & Company. Available at https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai.

Snidaro, L. (2023) ChatGPT act as an intelligence officer. Proceedings of the 2023 IEEE International Workshop

on Technologies for Defense and Security (TechDefense), 449–454.
Snyder, H. (2019) Literature review as a research methodology: An overview and guidelines. Journal of Business Research, Vol. 104, 333–339.
Strasser, A. (2023) On pitfalls (and advantages) of sophisticated large language models. arXiv preprint, arXiv:2303.17511.
Tang, A., Li, K. K., Kwok, K. O., Cao, L., Luong, S. and Tam, W. (2023) The importance of transparency: Declaring the use of generative artificial intelligence (AI) in academic writing. Journal of Nursing Scholarship, Vol. 56, No. 2, 314–318.
Thormundsson, B. (2024) Artificial intelligence (AI) worldwide — Statistics & facts. Statista. Available at https://www.statista.com/topics/3104/artificial-intelligence-ai-worldwide/.
Titko, J., Steinbergs, K., Achieng, M. and Uzule, K. (2023) Artificial intelligence for education and research: Pilot study on perception of academic staff. Virtual Economics, Vol. 6, No. 3, 7–19.
Tranfield, D., Denyer, D. and Smart, P. (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, Vol. 14, No. 3, 207–222.
Tu, X., Zou, J., Su, W. J. and Zhang, L. (2023) What should data science education do with large language models. arXiv preprint, arXiv:2307.02792.
Upshall, M. (2019) Using AI to solve business problems in scholarly publishing. Insights, Vol. 32, No. 1.
Webster, J. and Watson, R. T. (2002) Analysing the past to prepare for the future: Writing a literature review. MIS Quarterly, Vol. 26, No. 2, xiii–xxiii.
Williams, D. O. and Fadda, E. (2023) Can ChatGPT pass glycobiology? Glycobiology, Vol. 33, No. 8, 606–614.
Zheng, Z., Ning, K., Chen, J., Wang, Y., Chen, W., Guo, L., et al. (2023) Towards an understanding of large language models in software engineering tasks. arXiv preprint, arXiv:2308.11396.
Zhu, Z., Li, B., Xu, Y. and Rudzicz, F. (2021) What do writing features tell us about AI papers? arXiv preprint, arXiv:2107.06310.