Are preprints a threat to the credibility and quality of artificial intelligence literature in the ChatGPT era? A scoping review and qualitative study

Abstract

ChatGPT, as the pioneer of advanced generative AI tools, has triggered scholarly discussions about the potential use of such AI technologies in interdisciplinary fields. With a focus on the surge of AI-related preprints since the introduction of ChatGPT by OpenAI, the study investigated what the surge implies for AI literature, particularly in terms of credibility and quality. A scoping review was initially conducted to study the characteristics of the AI-related preprints in the Web of Science (WoS) database and also in five (5) preprint platforms (ArXiv, MedRxiv, SocArxiv, SSRN, and Research Square). The publication date range was set at 2023-01-01–2023-08-09. This was followed up by an interpretive phenomenological analysis (IPA) of the perceptions of experts in the AI field about the preprints. Employing a scoping review of AI-related preprints across six (6) databases and a qualitative analysis of fifteen (15) AI experts' opinions, our study reveals concerns about the research accuracy, quality, and credibility of preprints, and advocates for a robust evaluation and high-quality assurance process to promote open science objectives during their dissemination. Specifically, 45,918 AI-related preprints were found in the 6 preprint databases or repositories across different fields. The nine (9) themes from the IPA showed that preprints can be of value. However, experts advocated for the safe and responsible use of AI-related preprints, involving such tenets as maintaining ethical integrity and high-quality work on the part of authors and establishing sound AI-content guidelines from publishers and editors. Future studies are recommended to investigate the impact of preprints on decision-making processes in educational research and practice.

1. Introduction

In recent years, there has been a surge of preprints in the field of artificial intelligence (AI). Since ChatGPT's introduction by OpenAI on November 30, 2022, the AI field has seen an unprecedented surge in preprints, raising critical questions about their impact on the field's credibility and research quality. When the World Wide Web was introduced, there was great apprehension relating to widespread low-quality scientific research that had not been put through the peer review process, which traditionally acted as a filter for journals (Smart, 2022). The proliferation of AI preprints in diverse disciplines, especially on large language models (LLMs) or conversational agents such as ChatGPT, has led to a debate on whether such preprints should be relied upon, especially preprints in the health sector or for clinical practice (Gianola et al., 2020). The ease of uploading preprints to disseminate research results has been misused to propagate bold claims without sufficient supporting evidence, leading to concerns about the studies' credibility and quality.

According to Sarabipour et al. (2019), "Preprints are online, freely available (open-access) scientific manuscripts posted by authors on dedicated servers prior to peer review and publication in an academic journal" (p.2). The COPE Council (2018) also defines preprints as "a scholarly manuscript posted by the author(s) in an openly accessible platform, usually before or in parallel with the peer review process" (p.1). Preprints, which are early versions of manuscripts shared online, have become a term that encompasses a range of different materials and resources. Unfortunately, this term is sometimes used to refer to post-print manuscripts as well as materials that were never meant to be published (Pourret & Ibarra, 2023).

The deluge of ChatGPT/AI-related preprints is partially due to the transformative impact ChatGPT and other LLMs such as Google Bard, Microsoft Bing Chat, and Anthropic Claude+ are

touted to bring to the educational arena and other significant sectors of society (Mijwil et al., 2023; Tlili et al., 2023; Van Wyk et al., 2023). ChatGPT, which is perhaps the pioneer of advanced AI-powered chatbots or conversational agents, garnered over a million users in the first month of its release because of its powerful features. The "chat" refers to a conversational attribute while the "GPT" stands for Generative Pre-trained Transformer (a machine learning model that understands and generates human-like language). The powerful features of ChatGPT include its ability to admit its own mistakes (Adarkwah, Amponsah, et al., 2023; Tlili et al., 2023), pass medical tests (Kung et al., 2023), compose music, write essays, song lyrics, and teleplays, as well as aid in research manuscript preparation (Lund & Wang, 2023). Thus, ChatGPT and its allied conversational agents represent a paradigm shift in the field of education and other areas of life (Adarkwah, Ying, et al., 2023; Tlili et al., 2023). It is therefore no surprise that ChatGPT/AI-conversational agents are a hot topic currently and that many preprints have been uploaded to provide a snapshot of its disruptive nature in education and other sectors.

Nonetheless, flawed studies (incomplete or with errors in the content or methods) that are not filtered by a rigorous review process might threaten the credibility and quality of AI literature. Although some preprint repositories have screening services, there is still a high possibility of the preprint report perpetuating misinformation because of the low-quality or nonexistent peer review process. The flood of data sets and working papers about AI conversational agents in this ChatGPT era might cause novice researchers in the AI field, interested readers, and even experts in the field to take published journal articles at face value as opposed to the preprints because the preprints are so easily accessible. Misinformation in AI can have a critically negative impact not only on educational research but also on practice. Designers, policymakers, and school stakeholders are informed about design solutions, policy initiatives, and school-decision, respectively, based on

research evidence. To date, few efforts have been made to investigate the danger ChatGPT/AI-related preprints might pose to AI literature.

Therefore, the current study aims to explore the potential threats posed by preprints to the credibility and quality of AI literature in the ChatGPT era. By analyzing preprints related to ChatGPT and AI and gathering input from field experts, we provide new knowledge on the rapid expansion of preprint documents and gain insights into the benefits and drawbacks of preprints and their impact on the quality of AI research. Our findings will contribute to ongoing discussions about the role of preprints in scientific communication and their potential impact on research quality. Scoping reviews are used to present an overview of a specific subject or field. Presently, to the best of our knowledge, no study or review has been conducted on preprints related to ChatGPT and their potentially significant impact on AI literature, and therefore this is an original contribution to this increasingly complex field.

To accomplish the aim of the study, we first sought to establish evidence of the rapid proliferation of preprints. This is captured in research questions one and two (RQ1 and RQ2). Research question three (RQ3), which is split into two sub-questions, focuses on the broad aim of the study, which seeks to solicit expert opinions on the potential implications of the preprints on the credibility and quality of AI literature.

RQ1. What is the prevalence of preprints in the AI literature in the ChatGPT era?

RQ2. What are the characteristics of preprints from the WoS database in the ChatGPT era in terms of region, author counts, citation count, and publication sources/titles?

RQ3. What are the attitudes of expert researchers towards preprints in the AI field?

      a) What are the perceptions of experts about the danger of ChatGPT-related preprints to AI literature and their potential implications across multidisciplinary fields?

b) What measures can be taken to ensure the reliability and accuracy of preprints in the AI field?

The paper is organized in the following manner: Section 2 provides a literature review on the topic. Specifically, 2.1.1 is a literature review on the history, benefits, challenges, and ethical issues relating to preprints. Section 2.1.2 is a literature review on the potential role of preprints in open science. Section 3 presents the methodology of the paper in terms of both the literature and qualitative analysis. Section 4 is a discussion provided to triangulate the findings of the literature and qualitative phase of the study. Lastly, Section 5 provides a summary, implications, and future research directions.

2.1 Literature review

2.1.1 Preprints: history, benefits, challenges, and ethics

Preprints are not a recent development or invention (Ginsparg, 2016) but have existed since the 1960s (Pourret & Ibarra, 2023; Smart, 2022). ArXiv.org, which was established by Paul Ginsberg in 1991 at Los Alamos National Laboratory in the USA, is one of the earliest official repositories for preprints (Smart, 2022). The launch of the ArXiv preprint platform precipitated other sharing initiatives between scientific researchers, particularly in the field of physics, as a means to make good use of the new World Wide Web. The platform has evolved over time to include several subject areas such as economics and mathematics. Smart (2022) also states that the Social Science Research Network (SSRN), established by a pair of financial economists, was the second preprint platform to be launched in 1994, hosting over a million articles. The SSRN platform was subsequently purchased by Elsevier in 2016, although the purchase by a commercial publisher was heavily criticized by the research community. Many preprint repositories have been launched since

2010, such as the Center for Open Science (COS) community preprint platform. Additionally, the COVID-19 pandemic accelerated the publishing of preprints, as it became a standard process for studying COVID-19 to bring about a change in the rate of infections (Smart, 2022). A research report showed that more than 30,000 articles on COVID-19 in 2020 were preprints (Else, 2020). In our current era, there are several regional preprint repositories such as ArabiXiv, AfricArxiv and IndiaRxiv (Mallapaty, 2020).

Preprints have several benefits, as they can aid in quickly attaining data insights and serve as an important means of disseminating research findings. As mentioned before, the sharing of research information during the COVID-19 pandemic (Smart, 2022) helped in monitoring the infection situation in many countries and rolling out health protocols needed to curb the rate of infection. For example, life-saving discoveries were made through preprint literature during the COVID-19 pandemic (Watson, 2022). In discussing the advantages of preprints, Fry et al. (2019) list three categories: credit, visibility, and review. First, publishing citable preprints helps researchers to prioritize their work using a public record. This practice has been accepted by funding bodies. Also, preprints are more visible to other researchers because of their open-access nature. Findings from researchers can be accessed more quickly than in traditional publication outlets. Lastly, preprints can complement the review process by allowing researchers to contact authors to request suggestions for improvement on yet-to-be-published work. Smart (2022) also enumerates a few benefits of preprints, including granting authors greater control by giving them a platform to post their research to without relying on editors to accept their work. Also, in some disciplines, it might be important to know a published article's history and how it developed following different iterations. In instances where there is bias in journal publications and

discrimination regarding which articles are published, preprints help in promoting a democratic publishing process.

Nonetheless, recent studies highlight the double-edged nature of preprints in AI and related fields: while they accelerate knowledge dissemination, they also pose challenges in ensuring research quality and credibility (Moshontz et al., 2021; Rieger, 2020; Xu et al., 2021). In discussing best practices in using preprints, Rieger (2020) states that we should be cautious of the potential negative consequences of technology when creating, validating, sharing, and archiving knowledge. One challenge of preprints is the worry that one's work could be copied before publication. However, it is uncertain whether this risk is greater than in comparable situations that can occur during peer review (COPE Council, 2018). According to Mallapaty (2020), the long-term survival of preprint repositories is under threat in emerging economies because of the cost of keeping the platforms running. This means that in such regions, important sources of research information might be lost if sufficient money is not raised to sustain the hosting service. For example, Mallapaty mentioned that preprint repositories such as INA-Rxiv and IndiaRxiv, which contribute to regional science, face the possibility of closure because of money issues. Another challenge the COPE council mentions is that the license agreement of preprint platforms may conflict with that of a journal in which a researcher wants to publish their manuscript.

Also, the COPE council emphasizes that there is a tendency for inappropriate screening and an inability to remove research material from a preprint platform when a major flaw is identified in the research report. Above all, many researchers hold the view that preprints have a negative impact on the credibility and perception of scientific research due to the lack of a rigorous review process. Particularly in the health sector, the impact of preprints can prove to be substantial when they are accepted as established evidence. To this end, Smart (2022) opines that many editors,

journals, and publishers have viewed preprints with suspicion due to a lack of scientific integrity, scooping or stealing of data, copyright problems, and key concerns about research ideas. For example, a 2020 survey reported that out of 383 Asian academic society journals from 22 countries, only 28 accepted submissions that had already been posted on preprint servers and only 8 journals allowed for preprints to be cited in the reference list. At the same time, 118 editors in Korea disagreed with the view that preprints are essential (Smart, 2022).

As an ethical issue, not all preprints are eventually published in accredited outlets or peer-reviewed journals (Smart, 2022). A research article in 2021 reported that preprints represent only 4% of published literature, 60% are never published, and 30% of the literature on a preprint platform called bioRxiv remains unpublished (Anderson, 2020; Smart, 2022; Xie et al., 2021). These statistics prompt the question: are preprints worth it? Another ethical concern is the fear that better work might be ignored while preprints (weak or unreviewed work) will be exaggerated or more visible in the media (Fry et al., 2019). That is, the ethical question to ask is whether preprints are good for science but bad for the public (Fox, 2018). In their preprint report, the COPE Council (2018) highlighted a myriad of ethical issues regarding preprints by posing the following five questions:

- Are preprints publications?
- Do preprints establish precedence?
- What happens to the preprint if the work is subsequently published in a journal?
- Can papers be posted on multiple preprint platforms?
- What are the license implications of posting on a preprint platform?

2.1.2 Preprints and open science

To promote open science practices, the Center for Open Science (COS) has developed a web infrastructure for preprint archiving services. The Open Science Framework (OSF) Preprints have three new services (PsyArXiv, SocArXiv, and engrXiv) as of December 2016 to offer free, open access, and open-source archives for psychology, social, and engineering science research. As of August 16, 2023, there were 2,404,519 searchable preprints in the COS platform. The OSF platform helps researchers obtain feedback from colleagues in their own field of research before publication. More than one million preprints on the platform can be accessed by subject category using specific search keywords. Kirkham et al. (2020) reports that to advance open science practices, reputable publishers such as Elsevier, Springer, and Wiley have developed, codeveloped, or acquired preprint platforms or services. Also, several funders and research institutions, such as the Wellcome Trust, have been provided with examples of F1000 research's post-publication peer review and publishing platform to promote preprint-first research publishing. Kirkham and colleagues add that SciELO in April 2020 introduced a preprint platform that works in conjunction with Open Journal Systems. Preprints support open science in diverse ways, such as publishing under open licenses while at the same time allowing for different versions of data sets, working papers, and other documents to be posted. From the foregoing information, it can be argued that preprints serve as a pathway to open science.

Figure 1 displays the number of new submissions received each month since August 1991 (a period of 32.1 years). The total number of submissions excludes 2,431 articles that were migrated to arXiv rather than being submitted directly and 155 articles that have been deleted. The total number of articles available is 2,311,756. Figure 2 shows the number of preprint submissions in ArXiv by category since 2019.
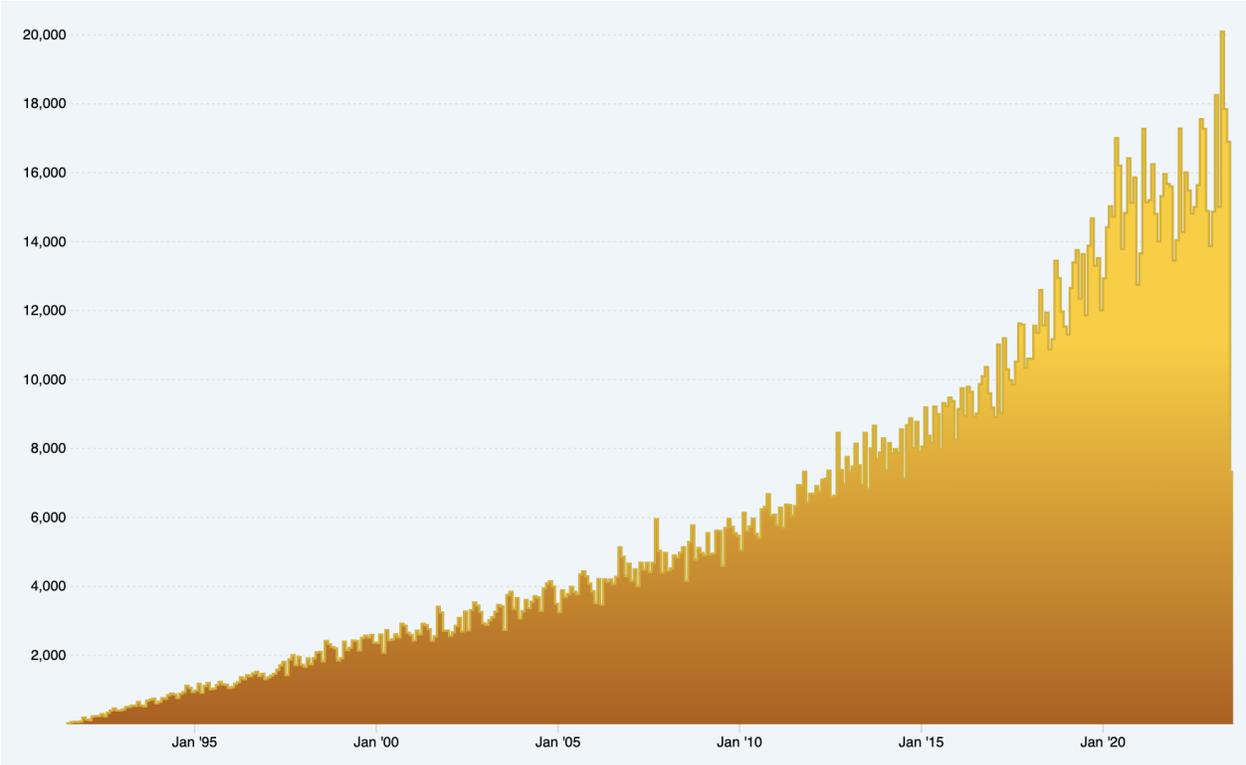
Figure 1 The total number of preprint submissions to ArXiv as of August 15, 2023 = 2,309,325.
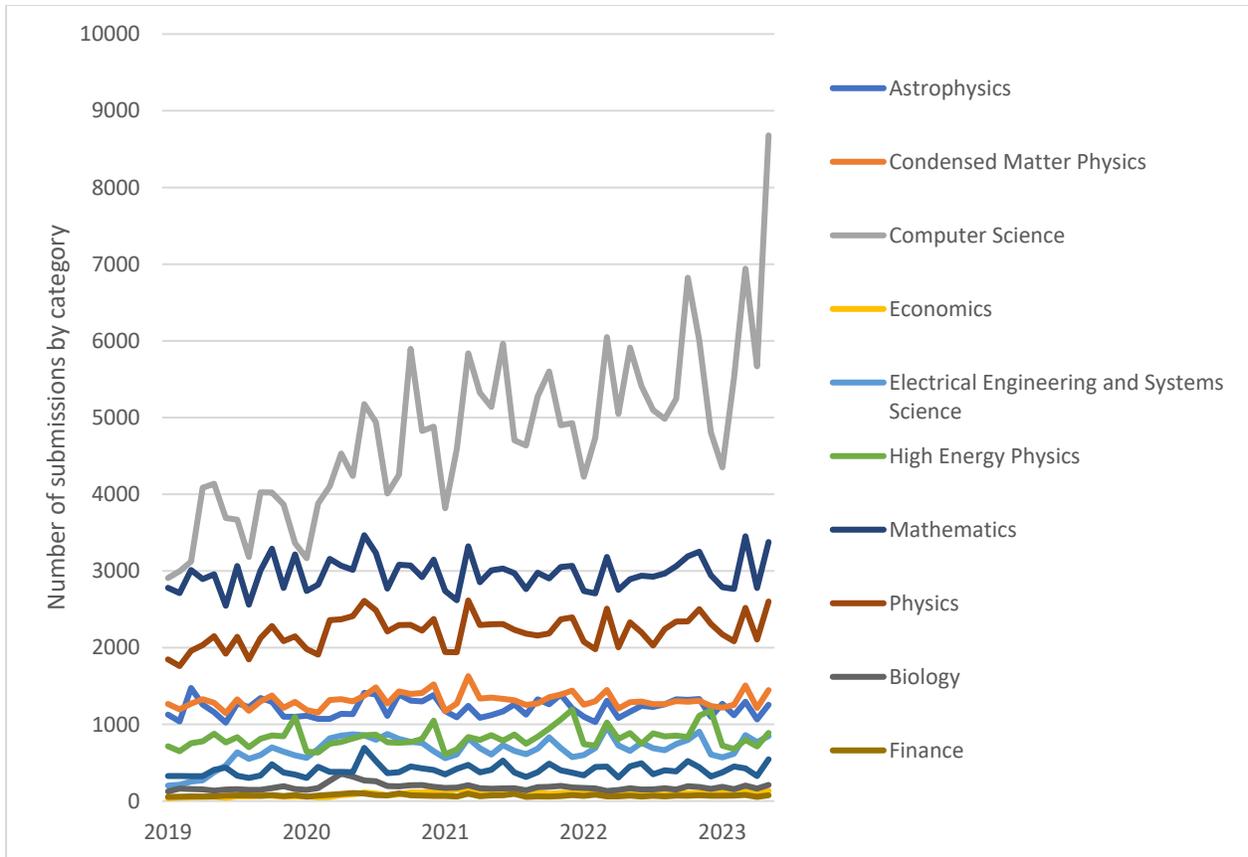
Figure 2 Number of preprint submissions in ArXiv by category since 2019

3. Method

In this study data triangulation was employed to gather data from literature and experts in the field of educational technology. Data triangulation is the use of multiple sources of data in the same study to validate research findings or enhance the robustness of data (Hussein, 2009). Triangulation has been used in the social sciences to enhance the credibility of research findings (Cheng et al., 2018; Hussein, 2009). The first step involved searching the Web of Science (WoS) database for preprints about ChatGPT as well as popular preprint repositories such as ArXiv, MedRxiv, SocArxiv, SSRN, and Research Square. The search was carried out on August 9, 2023. The final step of the study involved in-depth semi-structured interviews with educational technology experts who had published papers on ChatGPT since its launch.

3.1 Scoping review

3.1.1 Search strategy and results

A search was carried out on the Web of Science (WoS) using the Preprint Citation Index database with the keywords "ChatGPT OR Generative Pre-trained Transformer OR GPT OR Large language models OR Natural language processing OR AI conversational agents OR chatbots". 2,322 preprints were obtained and documented. The same keyword was used in the ArXiv and MedRxiv preprint repositories and yielded 1,838 results and 7,439, respectively. Using the keyword "ChatGPT" in SocArxiv, 13,339 results were generated. A search using the keyword "ChatGPT" was also conducted in SSRN using title, abstract, keywords, and full-text, and it generated 260 preprints. The final search was conducted on Research Square using the keywords "ChatGPT OR GPT OR Large Language Models" and yielded 20,720 preprints. In all, 45,918 preprints were found across the six (6) preprint databases or repositories (see Table 1).

Table 1 Database/repositories and record count (2023-01-01–2023-08-09)

No.	Preprint database/repository	Number of preprints
1.	WoS Preprint Citation Index	2,322
2.	ArXiv	1,838
3	MedRxiv	7,439
4.	SocArxiv	13,339
5.	SSRN	260
6.	Research Square	20,720

The publication date range was set at 2023-01-01–2023-08-09. There was no restriction on the language, discipline or research area of the articles or method used. However, all the documents from the WoS Preprint Citation Index database were in English. From WoS, the 2,322 preprint documents had been cited 525 times without self-citations from 330 citing articles, with 0.23 average citations per item and a 10 H-index (see Figure 3).
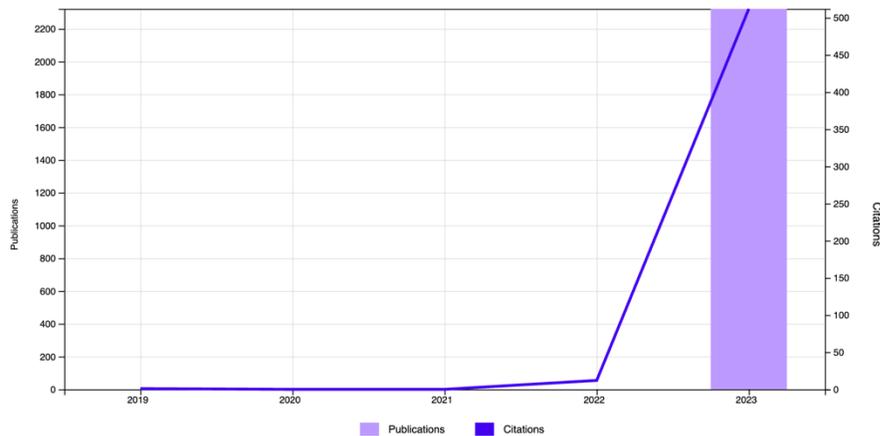
Figure 3 Citation report of records from WoS Preprint Citation Index

The publication sources/titles of the 2,322 articles are presented in Table 2. With 69 documents, Google Research is the top leading organization in ChatGPT research from the WoS Preprint Citation Index, followed by Microsoft research, with 53 documents, then Stanford University, with 51 documents (see Figure 4).

Table 2 Publication sources/titles and record count (2023-01-01–2023-08-09)

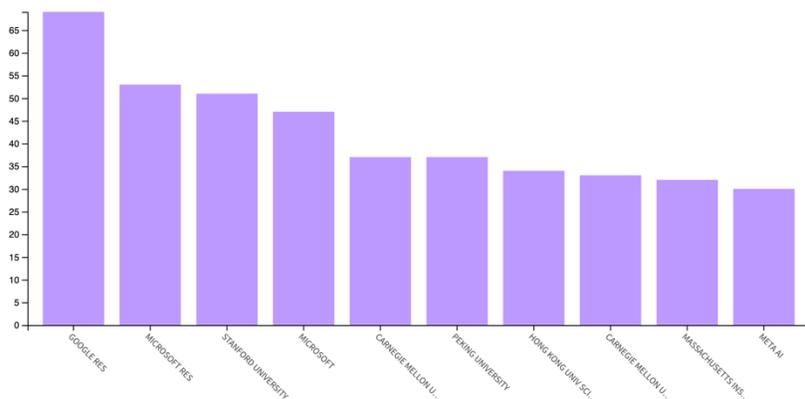| No.	Publication source/titles	Record count	% of 2,322
	ArXiv	2,144	92.334%
	MedRxiv	70	3.015%
	BioRxiv	65	2.799%
	Preprints	29	1.249%
	ChemRxiv	14	0.603%

Figure 4 Leading organizations in ChatGPT research from WoS Preprint Citation Index

According to the WoS Preprint Citation Index, the majority of authors involved in ChatGPT research are based in the United States, from which 1,145 documents originate (see Figure 5). In this study, ChatGPT-related preprints and AI-related preprints are used interchangeably. AI-related preprints mean preprints in this ChatGPT era.

Figure 5 Geographical region of top authors in ChatGPT research from WoS Preprint Citation Index

China is the second most prevalent country of origin for ChatGPT research, with 508 documents, followed by England, with 193 documents. Table 3 is a list of the top 10 most cited ChatGPT-related papers in the WoS Preprint Citation Index as of 2023, along with their authors, affiliations, paper titles, disciplines, databases, and citation counts.

Table 3 Information of top 10 cited authors from WoS Preprint Citation Index (2023-01-01–2023-08-09)

No.	Author(s)	Affiliation	Paper title	Discipline(s)	Database	Citation count
1.	OpenAI (2023)	OpenAI	GPT-4 Technical Report	Computer Science, Interdisciplinary Applications, Artificial Intelligence	ArXiv	35
2.	Bubeck et al. (2023)	Microsoft Research	Sparks of Artificial General Intelligence: Early Experiments with GPT-4	Computer Science, Interdisciplinary Applications	ArXiv	28
3.	Zhuo et al. (2023)	Monash University	Exploring AI Ethics of ChatGPT: A Diagnostic Analysis	Computer Science, Interdisciplinary Applications, Software Engineering	ArXiv	14
4.	Touvron et al. (2023)	Meta AI	LLaMA: Open and Efficient Foundation Language Models	Computer Science, Interdisciplinary Applications	ArXiv	13
5.	Eloundou et al. (2023)	OpenAI	GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models	Economics, Computer Science, Artificial Intelligence	ArXiv	11
6.	Kosinski (2023)	Stanford University	Theory of Mind May Have Spontaneously Emerged in Large Language Models	Computer Science, Interdisciplinary Applications	ArXiv	11
7.	Bang et al. (2023)	Hong Kong Univ Sci & Technol, Ctr Artificial Intelligence Res CAiRE	A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity	Computer Science, Interdisciplinary Applications, Artificial Intelligence	ArXiv	11
8.	Guo et al. (2023)	Shanghai Univ Finance & Econ, Sch Informat Management & Engn, AI Lab	How Close is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection	Computer Science, Interdisciplinary Applications	ArXiv	11
9.	Borji (2023)	Quintic AI, San Francisco	A Categorical Archive of ChatGPT Failures	Computer Science, Interdisciplinary Applications; Artificial Intelligence	ArXiv	10
10.	Sallam (2023)	Univ Jordan, Sch Med, Dept Pathol Microbiol & Forens Med	The Utility of ChatGPT as an Example of Large Language Models in Healthcare Education, Research and Practice: Systematic Review on the Future Perspectives and Potential Limitations	Health Care Sciences & Services	MedRxiv	10

The papers cover a range of topics, a majority of which are in the field of computer science and interdisciplinary applications, with some also in economics and the healthcare sciences. Almost all the papers were from the ArXiv preprint repository. The most cited paper (35 citations) was the technical report from OpenAI, the developer of ChatGPT.

3.2 Qualitative study

After establishing evidence of the rapid dissemination of AI-related preprints in the ChatGPT era through the literature analysis, the qualitative phase of the study probed the potential implications of the preprints on the credibility and quality of AI literature. The qualitative method is discussed in detail below.

3.2.1 Research design

Interpretive phenomenological analysis (IPA), which holds that "understanding how individuals make sense of their experiences is an interpretative activity best accomplished through the detailed examination of particular cases within phenomena of interest" (Miller et al., 2018, p. 242), was used to gain expert opinions about preprints in the AI field in the ChatGPT era. In this study, the case under exploration was ChatGPT-related preprints in the AI field. We sought to gain an understanding of experts' attitudes towards preprints in the AI field in the ChatGPT era. IPA can help in exploring these perceptions and experiences in a nuanced and detailed way, providing insights into the complex factors that shape the use and dissemination of preprints in AI research. The overarching research question specifically focused on two sub-questions:

1. What are the perceptions of experts about the benefits and danger of ChatGPT-related preprints to AI literature and their potential implications across multidisciplinary fields?

2. What measures can be taken to ensure the reliability and accuracy of preprints in the AI field?

3.2.2 Participants and data collection

A convenience sampling approach was used to recruit professors, Heads of Departments, lecturers, and researchers in the field of AI or educational technology to provide answers to the research questions. Other criteria for selecting the participants were their familiarity with ChatGPT and their publication record. Each of the participants had to have published at least one paper on AI or generative AI or be familiar with ChatGPT. Their ratings regarding familiarity with ChatGPT on a scale of 1-5 (where 1 was not familiar and 5 was very familiar) showed an average rating of 4, which signified participants' appropriateness for the study. Diversity was ensured in recruiting the participants. Both male and female scholars were included in the study. There were a total number of fifteen (15) participants (6 females and 9 males) across different research foci in the AI field. In total, the participants have published 95 AI-related studies since 2022. The different research foci of the participants allow for different perspectives about the phenomenon under study. Participants were recruited from four continents (USA, Europe, Africa, and Asia) to allow for generalization and inclusion. That is, experts were selected based on their extensive experience in AI and educational technology, ensuring a diverse representation of perspectives from various sub-fields and geographical regions. Specifically, experts came from the UK, USA, China, Egypt, and Indonesia. The diverse academic disciplines of the experts include AI and Educational or Educational Technology, Technology-Enhanced Learning, Explainable AI and Natural Language Understanding or Processing, Learning Analytics, Machine Learning, Human-Computer Interaction, Character Computing, Intelligent Tutoring Systems, and so on. As can be observed, the selection of representative experts from different continents or countries helps to give context to the unique situation of the impact of generative AI tools such as ChatGPT on education, particularly, scientific (AI) literature. Additionally, the various academic disciplines provide a

strong foundation in different aspects of AI and education, which qualifies the experts to draw on their experiences and knowledge to predict the potential implications of ChatGPT-related preprints on the credibility and quality of AI literature. Table 4 presents a summary of participants' personal information.

An introductory letter was sent to each participant to inform them about the aims of the study and the potential benefits of their participation in the study for the education and research community. After we had gained consent to collect data, the participants were given access to the interview guide in the form of a questionnaire through which we would collect their feedback. This was appropriate for the study because of the busy schedule and varying time zones of the participants involved. Data collection began on October 1, 2023, and ended on December 10, 2023.

Table 4 Personal information of participants

Participant	Age	Gender	Country	Position	Research Focus	AI papers published since 2022	Familiarity with chatbots
P1	65	Female	UK	Full Professor	AI and Education	20	5
P2	78	Male	USA	Regents Professor	AI and Educational Technology	10	5
P3	27	Male	China	Young Professor	Technology-enhanced learning	2	4
P4	45	Female	Egypt	Associate Professor	Explainable AI, Natural Language Understanding	8	4
P5	36	Male	China	Assistant Professor	Creativity & AI	0	2
P6	43	Male	Indonesia	Researcher	Engineering	0	4
P7	34	Female	China	Associate Researcher	Learning Analytics	4	3
P8	35	Male	Indonesia	Lecturer	Machine Learning, Smart City	2	3
P9	30	Female	China	Assistant Professor	Informal online learning, AI-empowering education	0	4
P10	38	Male	China	Associate Professor	Intelligent Tutoring Systems	10	5
P11	40	Male	China	Associate Professor	Learning analytics & AI	0	3

P12	31	Female	Egypt	Lecturer	Neuro-Symbolic Artificial Intelligence	6	5
P13	61	Male	USA	Full Professor	Intelligent Tutoring Systems	20	4
P 14	63	Male	Egypt	Professor	Machine Learning	4	4
P15	34	Female	Egypt	Assistant Professor	Applied Machine Learning, AI, Human Computer Interaction, Privacy Preservation, Character Computing	9	5

3.2.3 Interview guide development

Two main techniques were used in constructing the interview guide. Firstly, two (2) of the authors separately constructed a list of questions from the literature reviewed. Secondly, a brainstorming section was organized between three (3) of the authors to reduce twenty (20) items initially developed into a set of nine (9) items. Some of the questions removed were duplicates or overlapping items. In the brainstorming section, some of the questions were organized into probes. The final list of 9 questions were considered appropriate for the study based on consensus among the 3 researchers. The interview guide was semi-structured with open ended questions to gain deep insight about participants' perceptions and experiences. Some of the questions included, "What do you think of the proliferation of ChatGPT-related preprints in general in scientific/AI research?", "What do you think are some benefits of the proliferation of ChatGPT-related preprints for AI literature?", and "In your opinion, what are the potential dangers of ChatGPT-related preprints to AI literature, and how can these dangers impact multidisciplinary fields?"

3.2.4 Ethical considerations

The participants were assured that their personal information would remain confidential and would not be disclosed or used in the analysis and reporting of the study's results. All the ethical principles

in the Helsinki Declaration were complied with in gathering, storing, and analyzing participants' data.

3.2.5 Data analysis

For the analysis of the gathered interviews, a content analysis approach was employed. Content analysis is a well-established method for examining textual materials (Flick, 2009). The analysis followed the steps outlined by Erlingsson and Brysiewicz (2017). Intercoder agreement was ensured between three coders. The generated codes guided the formulation of the themes of the study. Table 5 presents some representative excerpts from responses under each of the nine (9) themes that emerged in the study.

Table 5 Definition of themes

Theme	Explanation
Preprints as publications	Use this theme when participants are talking about their general perception of ChatGPT-related preprints and their consideration as publications
Benefits of ChatGPT-related preprints	Use this theme when participants are talking about the positive side of ChatGPT-related preprints
Potential dangers of ChatGPT-related preprints	Use this theme when participants are talking about the negative side of ChatGPT
ChatGPT-related preprints and open science	Use this theme when participants are talking about ChatGPT-related preprints together with open science
Citing ChatGPT-related preprints	Use this theme when participants are talking about including or citing ChatGPT-related preprints in their manuscript
Credibility of AI literature	Use this theme when participants are talking about what ChatGPT-related preprints mean for the credibility of AI literature
Ethical issues of AI-related preprints	Use this theme when participants are talking about ethical issues of AI-related preprints in the ChatGPT era.
ChatGPT-related preprints and peer review process	Use this theme when participants are talking about ChatGPT-related preprints and peer review process
Safe and Responsible use of AI preprints	Use this theme when participants are talking about measures to ensure safe and responsible use of AI literature in the ChatGPT era.

3.2.5 Findings

The qualitative analysis identified key themes such as concerns about research quality, the need for more rigorous peer review processes, and the potential for preprints to contribute positively to open science if certain criteria are met.

Preprints as publications

Many participants were concerned about the rapid proliferation of ChatGPT-related preprints in AI research. Generally, they advocated for a thorough assessment of the preprints before usage. Particularly, there were mixed views regarding the consideration of such preprints as research publications. While some participants were totally against considering the preprints as scientific publications in AI research, a segment of the participants mentioned they might be useful and appropriate if the preprints were from renowned authors, as shown below:

> I have some concerns about the speed at which ChatGPT-related research is being disseminated through preprints before going through a robust peer review process. While making cutting-edge work quickly available can accelerate progress, we must be cautious about making claims not yet thoroughly vetted (P1, Full Professor, UK).

> I believe the proliferation of ChatGPT-related preprints in the realm of scientific and AI research is an inevitable phenomenon… I would certainly regard ChatGPT-related preprints as legitimate publications and would cite them in my work. However, my decision to do so would be based on the academic merit of the paper itself, rather than its association with ChatGPT (P7, Associate Researcher, China).

They still represent a body of work available to the public that can be critiqued. If authored by renowned researchers, there is little reason to suspect the authenticity of their work (P4, Associate Professor, Egypt).

Benefits of ChatGPT-related preprints

There was a consensus among the participants that ChatGPT-related preprints had several benefits to some extent although some called for human monitoring or assessment. A participant mentioned that such preprints can help ascertain knowledge for successful human-machine collaboration. Additionally, a majority of participants also believed that researchers could conceptualize a new study after reading the preprint literature, as expressed below:

I see few if any significant and sustained benefits to a sizeable population ... if monitored by humans. (P2, Regents Professor, USA).

It is helping advance the field and avoid revisiting already covered terrain. Still, in my opinion, it shouldn't be regarded separately from normal preprints (P15, Assistant Professor, Egypt).

In my opinion, one of the benefits is that ChatGPT-related preprints can help researchers formulate their future research ideas (P8, Lecturer, Indonesia).

Potential dangers of ChatGPT-related preprints

Research quality was an integral issue when talking about the potential dangers of ChatGPT-related preprints in AI literature. The propensity for misapplication of the findings of the preprints, especially for novice researchers, was a core issue raised by the participants. Nonetheless, a minority of respondents suggest that since some preprints will eventually be transformed into journal articles, they could be useful, but only after critical evaluation. As can be seen below:

> Potential drawbacks of preprints include a lack of peer review, implying that the quality and accuracy of the research might be questionable. Moreover, preprints might be based on brief experimental durations, making their findings potentially less enduring. In multidisciplinary fields, there's a risk of knowledge misapplication, as researchers might not be well-acquainted with insights from other disciplines, making it challenging to discern the accuracy of preprints, thereby potentially leading to the misapplication of these research outcomes (P3, Young Professor, China).

> A potential threat lies in the varied quality of these preprints. Some preprints may eventually be transformed into journal articles, while others might not…This presents higher demands for scholars as it necessitates a critical evaluation of information sourced from preprints (P9, Assistant Professor, China).

> If quality and ethical issues can be addressed, the potential dangers could be reduced. Critical review from domain experts plays the role of watchdog for research quality, which is especially important for multidisciplinary study (P11, Associate Professor, China).

ChatGPT-related preprints and open science

Responses from a large body of the participants revealed that ChatGPT-related preprints align with the objectives of open science, such as accelerating accessibility of scientific knowledge and discovery. For example, it was mentioned that the ease of accessing the preprints creates room for an interdisciplinary dialogue and helps promote open science after careful assessment:

> In principle yes, although some norms may need revisiting given ChatGPT's pace and capacity. As a field we must reflect on how to uphold scientific ideals around transparency, vetting claims, and guiding high-quality work. With thoughtfulness, preprints could advance open science (P1, Full Professor, UK).

> Yes, ChatGPT-related preprints align with open science objectives by enabling rapid knowledge sharing, fostering public accessibility, and encouraging collaborative input. They enhance transparency in research and can stimulate interdisciplinary dialogue. However, they also come with challenges like quality control and ethical considerations (P13, Full Professor, USA).

> Yes. Methodologically. And no, I would not upload preprints (P5, Assistant Professor, China).

Citing ChatGPT-related preprints

A minority of participants were opposed to the inclusion of ChatGPT-related preprints in their manuscript preparation on the basis that they are not research publications. About 20% of the

participants who were journal editors revealed that they would not accept preprint citations in the journals they are associated with. However, an ample number of the participants agreed that they would cite the preprints after rigorous evaluation. One participant mentioned his last resort would be the preprint literature when he has no published literature to cite from. As indicated in the statements below:

> Yes. Some preprints include some interesting findings. I will make sure though I make it clear when citing them that these are only preliminary findings (P12, Lecturer, Egypt). Journals should accept preprint citations to include cutting-edge research and promote open science. However, they must exercise caution by requiring authors to clearly mark preprint citations and update them if peer-reviewed versions become available. This approach balances rapid knowledge dissemination with quality assurance and ethical integrity (P13, Full Professor, USA).

> Firstly, I don't believe that the topics I am interested in can only be cited from preprints. Haha. However, if that were the case, I would decide whether to cite preprints after carefully assessing the quality of the articles (P9, Assistant Professor, China).

Credibility of AI literature

Neglecting to fact-check findings of ChatGPT-related preprints was considered harmful for AI literature in general. Some of the potential harm involves the difficulty in accessing good literature as a consequence of the influx of low-quality preprint papers. Other potential harmful issues are

plagiarism, issues of public trust or confidence in AI research, misinformed decision-making, and the increased risk of disseminating inaccurate misleading, or flawed research as indicated below:

> Unfortunately, there is a risk of erosion in credibility, if unchecked. Therefore, leaders in AI research must promptly define norms and processes for navigating this new era responsibly—upholding quality, scientific ideals, and public trust. The stakes are high, so we must get this right (P1, Full Professor, UK).

> Yes. It can increase the difficulty of literature search, and scholars might easily get lost amidst a vast sea of low-quality papers. Concurrently, I've observed that the efficiency of using large language models for searching papers is much lower compared to using them for understanding and writing papers (P7, Associate Researcher, China)

Regardless, one participant indicated that the positive gains of the preprints outweigh the harm, as reported below:

> I cannot quantify the impact. The good for me is more than the bad (P4, Associate Professor, Egypt).

Ethical issues of AI-related preprints

Participants in the study raised several ethical issues involving the proliferation and usage of AI-related preprints in the ChatGPT era. Key among the ethical concerns is the diminishing of academic integrity, misleading findings or claims from the preprint literature, and the possibility of unauthorized use of information. Additionally, it was also mentioned data privacy (i.e., the use

of personal data that hasn't undergone ethical review) and questionable methods or conclusions as ethical concerns of AI-related preprints. As stated below:

> As far as I am concerned, preprints are not indexed in some plagiarism detection engines. As a result, plagiarizing content from preprints might go undetected during plagiarism checks, potentially leading to issues of preprints being plagiarized. Meanwhile, plagiarism checks for preprints are not stringent, making it difficult to detect when a preprint has plagiarized from other articles (P3, Young Professor, China).

> Ethical issues centre on inadvertently misleading peers and the public through unchecked claims. Research integrity processes must evolve to address risks in this fast-changing era. Also, equity issues if some voices dominate without safeguards. Community dialogue on ethics and values is vital (P1, Full Professor, UK).
> Plagiarism, misuse of unreliable information, and knowledge property violation (P11, Associate Professor, China).

ChatGPT-related preprints and peer review process

The main recommendation arising from the study for ensuring rigor in the use of preprints was a robust and sound peer review process. Participants in the study mentioned the need for establishing clear criteria for assessing the quality of the preprints and a research community review. Some participants encouraged all preprint platforms to create a medium for authors to update their manuscript. Additionally, it was suggested that novice researchers need to be coached on how to

make use of preprints. A participant also mentioned the need of using AI assessors, detectors, or applications such as ChatGPT in the peer review process. As demonstrated below:

> Preprints could help rapidly surface promising work, provided the field implements heightened expectations and systems for vetting claims through peer input before findings are treated as established. Standards, oversight, and mentoring of early career researchers will be key (P1, Full Professor, UK).

> If possible I hope the preprint platform such as arXiv could make use of generative AI such as ChatGPT to be the reviewers. ChatGPT may have the ability to review a work critically with appropriate prompts (P10, Associate Professor, China).

> I think perhaps we don't need to introduce traditional peer review into the preprint vetting process. Instead, we should consider establishing open peer evaluation metrics for preprints, similar to how GitHub operates as an open-source code community (P9, Assistant Professor, China).

Safe and Responsible use of AI preprints

Handling and utilizing the preprints in a safe and responsible manner was considered a responsibility of the author and also the research community involved in the review process. Researchers in particular should submit complete and high-quality studies. The need for an AI-content use guideline was considered paramount. Also, it was recommended by one of the participants that preprint platforms should label the preprints as not peer reviewed:

Establishing a clear AI-content use guideline and improving the review process by ensuring a collaboration between reviewer and AI content detector (P11, Associate Professor, China). I think that a fundamental quality of a scholar is to ensure the caliber of every article they produce to the best of their ability. Furthermore, when there's a need to cite preprints, it's essential to diligently assess the quality of the preprints, scrutinize the rigor of their research processes, and maintain a critical perspective towards statements provided within these preprints (P9, Assistant Professor, China).

I encourage exploring this paradigm responsibly - but not overstating findings without due input; sound signaling to media on appropriate cautions; and proactively addressing integrity concerns through guidelines suited to the pace of modern AI. If we lead conscientiously, we can unlock potential (P1, Full Professor, UK).

4. Discussion

Through a scoping review and qualitative analysis, the study explored the prevalence of AI preprints in the ChatGPT era and the implications of such preprints for education and research in the AI field. In doing so, we examined the characteristics of preprints from the WoS database and other preprint repositories (ArXiv, MedRxiv, SocArxiv, SSRN, and Research Square). A qualitative analysis of expert opinions on the prevalence of AI-related preprints was conducted to explore the attitudes of expert researchers towards preprints in the AI field. Our findings suggest a need for more stringent quality controls in the dissemination of AI preprints to preserve the field's credibility, alongside embracing their potential in advancing open science.

Addressing our first research question, we found a significant surge in the number of preprints related to AI in the ChatGPT era. This was the case when COVID-19 emerged (Blatch-Jones et al., 2023; Else, 2020; Smart, 2022). Since the introduction of ChatGPT by OpenAI, there has been an unprecedented increase in the sharing of preprints, reflecting the transformative impact and interest generated by large language models (LLMs) and conversational agents. This proliferation raises concerns about the reliance on AI preprints, particularly in critical sectors such as health and clinical practice. Disseminating low-quality scientific findings can be particularly detrimental, especially when the research can directly impact medical practices (Kwon, 2020). As shown in Figure 2, the AI-related preprints cut across different sectors such as health and computer science where practitioners can make a decision or implement an intervention based on a study's findings.

Moreover, a thorough examination was performed on the characteristics of preprints from different repositories and the WoS database. In examining the characteristics of preprints from the WoS database, we analyzed region, author counts, citation count, and publication sources/titles. This analysis provided insights into the global distribution of preprints and the collaborative nature of AI research. We observed a diverse range of authors contributing to preprints, indicating the multidisciplinary nature of AI research. As reviewed earlier, the rapid dissemination of preprints in sensitive areas or disciplines such as health and psychology can cause harm if such studies report falsified findings (e.g., the use of a drug or treatment approach that is unverified or approved). In the area of research methodology responsible for equipping novice and emerging scientific scholars with the knowledge and skills to design, analyze, and interpret research studies, producing findings about false research methods, models, or instruments for collecting data can increase the rate of retraction of published studies and compromise research quality. As reported by Ioannidis

(2005) and Kraemer (2015), there is an increasing concern that the majority of current research findings are inaccurate. That is, it is crucial to note that preprints may lack the rigorous peer review process associated with traditional journal publications, potentially impacting their credibility and quality (Blatch-Jones et al., 2023; Klebel et al., 2020; Kwon, 2020). In this regard, preprint repositories such as bioRxiv and medRxiv enhanced their usual screening procedures in sharing COVID-19 literature (Kwon, 2020). As a recommendation, it is crucial to follow the COPE guidelines to maintain the integrity and credibility of academic publications (Mijwil et al., 2023).

Furthermore, through interviews, we explored expert scholars' perceptions of the benefits and dangers posed by ChatGPT-related preprints to AI literature and their potential implications across multidisciplinary fields. The findings revealed a mixed perspective among experts. While some acknowledged the value of preprints for rapid dissemination of research findings, others expressed concerns about the potential for misinformation and low-quality research being propagated through preprints. The lack of a peer review process was identified as a key factor contributing to these concerns. Nonetheless, peer-reviewed articles might be retracted due to errors or unreliable findings (Blatch-Jones et al., 2023). Hence, the peer review process for both preprints and journal articles should be robust. Preliminary findings of studies should be presented in a credible manner through expert comments, discussion, critique, and feedback.

In the qualitative phase of the study, expert suggestions on ensuring the reliability and accuracy of preprints in the AI field were implementing stricter screening processes for preprint repositories, promoting transparency in reporting methods and results, encouraging collaboration and peer feedback through open commenting systems, and fostering a culture of critical evaluation among researchers and readers. Additional quality assurance practices that can help promote academic trust and integrity involve checking for redundant publications, guest or ghost authorship,

ensuring disclosure of funding sources and competing interests, misreporting of funder involvement, and failure to obtain informed consent (Cogan, 2022). It is also recommended for preprint repositories and journal editors to check scientific citations in the body of a submitted manuscript to ensure their accuracy. This is a core responsibility of preprint platforms, publishers, and editors. However, a recent study found that about 75% of journals have no clear guidelines on co-reviewing, citation of preprints and publication of reviewers' identities for preprints (Klebel et al., 2020).

5. Conclusion, Implications, and Recommendations

The study aimed to explore the potential threats posed by preprints to the credibility and quality of AI literature in the ChatGPT era. A scoping review across six (6) databases and a qualitative approach (IPA) involving fifteen (15) AI experts were adopted. The current investigation revealed that since the launch of ChatGPT and its allied technologies, there has been a surge in the number of preprints related to AI. A total of 2,322 preprint documents have been cited 525 times without self-citations from 330 citing articles from WoS. The majority of authors involved in ChatGPT research from the WoS Preprint Citation Index were based in the USA, China, and England. The preprint documents published from the 6 databases covered diverse topics in different fields such as computer science, economics, and healthcare sciences.

The IPA during the qualitative phase revealed that experts in the field of AI had a myriad of concerns about the proliferation of AI-related preprints in the field. Some of the key concerns experts had on the impact of ChatGPT-related preprints on AI literature involved research accuracy, quality, and credibility. Specifically, experts had mixed views about the consideration of the preprints as research publications. They perceived the preprints to be of some value but advocated

for human monitoring or assessment. Potential dangers of the preprints involved their misapplication by novice researchers, misleading or questionable findings, and disregard for research ethics, which threaten the quality and credibility of AI literature.

The findings of the study imply that without assessing or vetting evidence from preprint literature, there is a potential harm to AI literature in general even though preprints offer some benefits. Without a safe and responsible use of AI-related preprints, the quality of academic research in the field might be questionable. Consequently, education practice that thrives on research evidence might be negatively affected. Arising from these findings, we call for evidence-based research and ethically vetted studies in the field. Researchers, policymakers, and stakeholders in the AI field should consider the need for critical evaluation when using preprints as sources of information and emphasize the importance of promoting transparency and rigorous peer review processes. Sound AI policies should be designed and implemented through stakeholder dialogue to make use of preprints an open science practice while at the same time maintaining the credibility of AI literature. It is also advocated that journal publishers and editors establish policies for including or citing preprints in AI literature. Also, a watchdog mechanism can be installed to check authors who report falsified research findings or engage in malpractice. Preprint repositories that produce a high percentage of falsified studies should face closure or be subjected to stringent measures to hold them accountable for the quality of studies mounted on their platform.

In the future, researchers should focus on developing guidelines and best practices for assessing the credibility and quality of preprints in AI literature. Additionally, investigating the impact of preprints on decision-making processes in educational research and practice would provide valuable insights into their practical implications. In the future, the establishment of a quality assurance model or scientific/AI system that can detect AI-generated content will be

beneficial in maintaining the credibility and quality of preprints. For example, we suggest investigations into systems that detect commonly generated words by ChatGPT and its allied technologies. Researchers in the information sciences, computer sciences, and linguistics can conduct further investigations into how to develop algorithms or models to be used by preprint repositories and journal editors to detect AI-generated content. Enhancing the credibility and quality of preprints supports the open science movement by ensuring that disseminated literature fosters confidence and trust in scientific literature. A limitation of the study is that while it provides valuable insights, its scope is limited to specific preprint platforms and the perspectives of a select group of experts, which may not fully represent the broader AI research community. To address these limitations, we propose that future studies involve more experts from different parts of the world and a greater range of institutions for a similar investigation. Research questions that explore how preprints influence policy-making in AI or assess their impact in various AI subfields, such as machine learning or natural language processing, would be valuable foci for future studies.

References

Adarkwah, M. A., Amponsah, S., Wyk, M. M. van, Huang, R., Tlili, A., Shehata, B., Metwally, A. H. S., & Wang, H. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning and Teaching*, *6*(2), Article 2. https://doi.org/10.37074/jalt.2023.6.2.26

Adarkwah, M. A., Ying, C., Mustafa, M. Y., & Huang, R. (2023). Prediction of Learner Information-Seeking Behavior and Classroom Engagement in the Advent of ChatGPT. In C. Anutariya, D. Liu, Kinshuk, A. Tlili, J. Yang, & M. Chang (Eds.), *Smart Learning for*

A Sustainable Society (pp. 117–126). Springer Nature. https://doi.org/10.1007/978-981-99-5961-7_13

Anderson, K. R. (2020). bioRxiv: Trends and analysis of five years of preprints. *Learned Publishing*, *33*(2), 104–109. https://doi.org/10.1002/leap.1265

Blatch-Jones, A. J., Saucedo, A. R., & Giddins, B. (2023). The use and acceptability of preprints in health and social care settings: A scoping review. *PLOS ONE*, *18*(9), e0291627. https://doi.org/10.1371/journal.pone.0291627

Cheng, Y.-W., Sun, P.-C., & Chen, N.-S. (2018). The essential applications of educational robot: Requirement analysis from the perspectives of experts, researchers and instructors. *Computers & Education*, *126*, 399–416. https://doi.org/10.1016/j.compedu.2018.07.020

Cogan, E. (2022). Preventing fraud in biomedical research. *Frontiers in Cardiovascular Medicine*, *9*. https://doi.org/10.3389/fcvm.2022.932138

COPE Council. (2018). *COPE discussion document: Preprints*. COPE: Committee on Publication Ethics. https://publicationethics.org/resources/discussion-documents/preprints

Else, H. (2020). How a torrent of COVID science changed research publishing—In seven charts. *Nature*, *588*(7839), 553–553. https://doi.org/10.1038/d41586-020-03564-y

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, *7*(3), 93–99. https://doi.org/10.1016/j.afjem.2017.08.001

Flick, U. (2009). *An introduction to qualitative research* (4th ed., pp. xxi, 505). Sage Publications Ltd.

Fox, F. (2018). *the preprint dilemma: Good for science, bad for the public? A discussion paper for the scientific community*. https://www.sciencemediacentre.org/the-preprint-dilemma-good-for-science-bad-for-the-public-a-discussion-paper-for-the-scientific-community/

Fry, N. K., Marshall, H., & Mellins-Cohen, T. (2019). In praise of preprints. *Microbial Genomics*, *5*(4), e000259. https://doi.org/10.1099/mgen.0.000259

Gianola, S., Jesus, T. S., Bargeri, S., & Castellini, G. (2020). Characteristics of academic publications, preprints, and registered clinical trials on the COVID-19 pandemic. *PLOS ONE*, *15*(10), e0240123. https://doi.org/10.1371/journal.pone.0240123

Ginsparg, P. (2016). Preprint Déjà Vu. *The EMBO Journal*, *35*(24), 2620–2625. https://doi.org/10.15252/embj.201695531

Hussein, A. (2009). The use of Triangulation in Social Sciences Research: Can qualitative and quantitative methods be combined? *Journal of Comparative Social Work*, *4*(1), Article 1. https://doi.org/10.31265/jcsw.v4i1.48

Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLOS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Kirkham, J. J., Penfold, N. C., Murphy, F., Boutron, I., Ioannidis, J. P., Polka, J., & Moher, D. (2020). Systematic examination of preprint platforms for use in the medical and biomedical sciences setting. *BMJ Open*, *10*(12), e041849. https://doi.org/10.1136/bmjopen-2020-041849

Klebel, T., Reichmann, S., Polka, J., McDowell, G., Penfold, N., Hindle, S., & Ross-Hellauer, T. (2020). Peer review and preprint policies are unclear at most major journals. *PLOS ONE*, *15*(10), e0239518. https://doi.org/10.1371/journal.pone.0239518

Kraemer, H. C. (2015). A Source of False Findings in Published Research Studies: Adjusting for Covariates. *JAMA Psychiatry*, *72*(10), 961–962. https://doi.org/10.1001/jamapsychiatry.2015.1178

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., Leon, L. D., Elepaño, C., Madriaga, M.,
Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of
ChatGPT on USMLE: Potential for AI-assisted medical education using large language
models. *PLOS Digital Health*, *2*(2), e0000198.
https://doi.org/10.1371/journal.pdig.0000198

Kwon, D. (2020). How swamped preprint servers are blocking bad coronavirus research. *Nature*,
581(7807), 130–131. https://doi.org/10.1038/d41586-020-01394-6

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact
academia and libraries? *Library Hi Tech News*, *ahead-of-print*(ahead-of-print).
https://doi.org/10.1108/LHTN-01-2023-0009

Mallapaty, S. (2020). Popular preprint servers face closure because of money troubles. *Nature*,
578(7795), 349–349. https://doi.org/10.1038/d41586-020-00363-3

Mijwil, M. M., Hiran, K. K., Doshi, R., Dadhich, M., Al-Mistarehi, A.-H., & Bala, I. (2023).
ChatGPT and the Future of Academic Integrity in the Artificial Intelligence Era: A New
Frontier. *Al-Salam Journal for Engineering and Technology*, *2*(2), Article 2.
https://doi.org/10.55145/ajest.2023.02.02.015

Miller, R. M., Chan, C. D., & Farmer, L. B. (2018). Interpretative Phenomenological Analysis: A
Contemporary Qualitative Approach. *Counselor Education and Supervision*, *57*(4), 240–
254. https://doi.org/10.1002/ceas.12114

Moshontz, H., Binion, G., Walton, H., Brown, B. T., & Syed, M. (2021). A Guide to Posting and
Managing Preprints. *Advances in Methods and Practices in Psychological Science*, *4*(2),
25152459211019948. https://doi.org/10.1177/25152459211019948

Pourret, O., & Ibarra, D. E. (2023). *The rise of preprints in earth sciences* (12:561). F1000Research. https://doi.org/10.12688/f1000research.133612.2

Rieger, O. Y. (2020). *Preprints in the Spotlight: Establishing Best Practices, Building Trust*. ITHAKA S+R. https://www.jstor.org/stable/resrep49517

Smart, P. (2022). The evolution, benefits, and challenges of preprints and their interaction with journals. *Science Editing*, *9*(1), 79–84. https://doi.org/10.6087/kcse.269

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Van Wyk, M. M., Adarkwah, M., & Amponsah, S. (2023). Why All the Hype about ChatGPT? Academics' Views of a Chat-based Conversational Learning Strategy at an Open Distance e-Learning Institution. *Open Praxis*, *15*, 214–225. https://doi.org/10.55982/openpraxis.15.3.563

Watson, C. (2022). Rise of the preprint: How rapid data sharing during COVID-19 has changed science forever. *Nature Medicine*, *28*(1), 2–5. https://doi.org/10.1038/s41591-021-01654-6

Xie, B., Shen, Z., & Wang, K. (2021). *Is preprint the future of science? A thirty year journey of online preprint services* (arXiv:2102.09066). arXiv. https://doi.org/10.48550/arXiv.2102.09066

Xu, F., Ou, G., Ma, T., & Wang, X. (2021). The consistency of impact of preprints and their journal publications. *Journal of Informetrics*, *15*(2), 101153. https://doi.org/10.1016/j.joi.2021.101153