

Article

Generative AI in Education: Perspectives Through an Academic Lens

Iulian Întorsureanu *, Simona-Vasilica Oprea , Adela Bâra  and Dragoș Vespan 

Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, 010374 Bucharest, Romania; simona.oprea@csie.ase.ro (S.-V.O.); bara.adela@ie.ase.ro (A.B.); dragos.vespan@ie.ase.ro (D.V.)

* Correspondence: iulian.intorsureanu@ie.ase.ro

Abstract: In this paper, we investigated the role of generative AI in education in academic publications extracted from Web of Science (3506 records; 2019–2024). The proposed methodology included three main streams: (1) Monthly analysis trends; top-ranking research areas, keywords and universities; frequency of keywords over time; a keyword co-occurrence map; collaboration networks; and a Sankey diagram illustrating the relationship between AI-related terms, publication years and research areas; (2) Sentiment analysis using a custom list of words, VADER and TextBlob; (3) Topic modeling using Latent Dirichlet Allocation (LDA). Terms such as “artificial intelligence” and “generative artificial intelligence” were predominant, but they diverged and evolved over time. By 2024, AI applications had branched into specialized fields, including education and educational research, computer science, engineering, psychology, medical informatics, healthcare sciences, general medicine and surgery. The sentiment analysis reveals a growing optimism in academic publications regarding generative AI in education, with a steady increase in positive sentiment from 2023 to 2024, while maintaining a predominantly neutral tone. Five main topics were derived from AI applications in education, based on an analysis of the most relevant terms extracted by LDA: (1) Gen-AI’s impact in education and research; (2) ChatGPT as a tool for university students and teachers; (3) Large language models (LLMs) and prompting in computing education; (4) Applications of ChatGPT in patient education; (5) ChatGPT’s performance in medical examinations. The research identified several emerging topics: discipline-specific application of LLMs, multimodal gen-AI, personalized learning, AI as a peer or tutor and cross-cultural and multilingual tools aimed at developing culturally relevant educational content and supporting the teaching of lesser-known languages. Further, gamification with generative AI involves designing interactive storytelling and adaptive educational games to enhance engagement and hybrid human–AI classrooms explore co-teaching dynamics, teacher–student relationships and the impact on classroom authority.



Academic Editors: Marko Horvat and Tomislav Jaguš

Received: 23 January 2025

Revised: 28 February 2025

Accepted: 3 March 2025

Published: 6 March 2025

Citation: Întorsureanu, I.; Oprea, S.-V.; Bâra, A.; Vespan, D. Generative AI in Education: Perspectives Through an Academic Lens. *Electronics* **2025**, *14*, 1053. <https://doi.org/10.3390/electronics14051053>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: education; gen-AI; large language models; bibliometrics; technology-enhanced learning

1. Introduction

1.1. General Context, Motivation and Objectives

Recently, generative artificial intelligence (gen-AI) has been an area of intense development and rapid evolution. Based on previous developments in natural language processing (NLP), transformer models began appearing from 2017 [1]. These have evolved onto an

increasingly larger scale, reaching hundreds of millions of parameters in 2018–2019 up to OpenAI’s GPT-3 model, with 175 billion parameters, in 2020. The launch of ChatGPT in November 2022 as a chatbot interface for GPT-3 for the general public started a new wave of innovation that aimed to integrate LLM capabilities into applications for various domains [2]. GPT-3.5, introduced in late 2022, marked an improvement over GPT-3 with better conversational abilities and enhanced context management. It handled longer prompts more effectively and produced more natural, human-like text. However, it still exhibited some limitations, such as occasional factual inaccuracies and verbosity and struggled with complex reasoning or maintaining consistency in extended interactions. The release of GPT-4 in early 2024 represented a major milestone in the development of gen-AI. It introduced multimodal capabilities, allowing the model to process and understand not only text but also images, opening up new possibilities in creative applications and data analysis. Version 4o symbolizes a leap in AI by addressing critical challenges involving memory, real-time data integration and advanced contextual understanding [3].

OpenAI’s ChatGPT has gained broad awareness among both the general public and the research community, although an ever-increasing number of competitors, both closed- and open-source, have also emerged. The focus of this paper is on the underlying LLM technology and gen-AI technologies in general and on their intersection with education, without any vendor-related limitations.

Early on, it became apparent that LLM technology would have a great impact on education [4]. On the one hand, it enables an unprecedented level of adaptive learning on a large scale, with AI-powered tutors being able to give personalized guidance for the individual student. For teachers, it supports the automation of time-consuming tasks, like creating exam questions [5] or assignment grading [6]. On the other hand, LLM brings in significant risks for academic integrity [7,8]; LLMs are equally accessible for cheating—either by providing responses to assessment questions or by writing essays. The problem is compounded by the difficulty in reliably detecting AI-generated content by anti-plagiarism tools [9].

These recent evolutions have sparked strong interest in researching the various applications and the impact of gen-AI for education, with numerous scholarly papers being published on these topics. The findings of our research are important for several reasons. We are motivated to achieve the following:

- (1) To provide a structured overview of a growing knowledge base in order to help users to stay informed about key trends and emerging topics as well as to enable the adoption of current developments and best practices;
- (2) To identify significant gaps and future research opportunities in order to ensure a balanced approach to the field under consideration;
- (3) To highlight influential contributors, publications and institutions that are shaping the field, enabling collaborative networks of interdisciplinary researchers to be developed;
- (4) To assess the effectiveness and impact of innovative gen-AI applications in teaching and learning as well as to trace the evolution of pedagogical approaches that are validated by research experiments;
- (5) To inform the debate about appropriate ethical guidelines, which respond to challenges related to privacy, intellectual property, student data usage and potential biases in AI algorithms.

Bibliometric analysis has emerged as a powerful method for systematically analyzing vast amounts of scientific data and gaining valuable insights into research trends. Its popularity stems from its capacity to handle extensive datasets—often comprising hundreds or thousands of publications—and its reliance on objective metrics such as citation counts, keyword occurrences and publication frequencies. Thus, bibliometric analysis allows researchers to uncover emerging trends, map collaboration networks and explore the

intellectual structure of specific fields, offering a comprehensive overview of the existing literature [10].

Conducting a bibliometric research of publications on education and gen-AI presents a series of challenges: (1) The collection of high-quality data about relevant publications requires a careful selection of academic databases that cover a wide range of disciplines. (2) Choosing an appropriate set of search keywords from a variety of used terms so that the relevant papers are retrieved in a consistent manner. (3) Accessing data on the newest publications; there is an inherent delay before a submitted manuscript is published and subsequently indexed in electronic databases. (4) Excluding irrelevant or low-quality studies from the selection can be a time-consuming process. (5) Possible limitations of using only quantitative metrics for analysis, which may fail to capture the quality and relevance of published studies.

This paper aims to provide a detailed bibliometric analysis of published papers related to education and gen-AI/LLMs in order to identify the most prominent trends in the current research, as well as potential gaps that require more focus. Apart from the bibliometric analysis, it also focuses on sentiment analysis and topic modeling, with the aim of identifying latent themes. Given the technological novelty of gen-AI/LLMs, their disruptive character as well as their rapid development, this analysis contributes to understanding the impact this technology has had on education in the approximately two years of LLMs' public availability.

1.2. Controversies in the Field of Education and AI

Controversies surrounding education and gen-AI span a wide range of issues, reflecting deep-seated concerns about the implications of AI on learning environments, intellectual integrity, student–teacher relationships and long-term impacts on society. Gen-AI's ability to produce essays, solutions and content has raised significant concerns regarding academic integrity. Critics argue that students may increasingly rely on AI tools to complete assignments, potentially undermining authentic learning and contributing to a rise in plagiarism [11]. This has sparked ongoing debates on the measures educational institutions may implement to effectively address the issue of AI-generated academic dishonesty [12,13]. Gen-AI produces content that blurs the lines of authorship and intellectual property, raising questions about who owns AI-generated content [14]. In educational settings, this poses challenges in evaluating originality and in respecting intellectual property rights for students and educators [13].

There is also a fear that students might become overly dependent on gen-AI for information retrieval, writing and problem-solving. This reliance could potentially weaken critical thinking skills, creativity and the ability to independently analyze and synthesize information, as students might lean on AI rather than develop their own understanding and perspectives [15]. Furthermore, gen-AI models are trained on vast datasets that may contain biases or inaccuracies, which could lead to the propagation of biased or incorrect information. When used in educational settings, this could reinforce stereotypes or spread misinformation, impacting students' perspectives and understanding of complex issues.

The integration of gen-AI into education challenges traditional teaching methods [16], raising concerns about the role of teachers and the importance of personal interactions in learning. Some argue that an over-reliance on AI may devalue the role of teachers, reduce classroom engagement and impact the quality of student–teacher relationships [17]. Additionally, critics argue that an over-reliance on gen-AI might not prepare students for future job markets that demand uniquely human skills like emotional intelligence, empathy and complex problem-solving [15]. By focusing heavily on AI tools, students may miss

out on developing interpersonal and analytical skills essential for career readiness in an environment where automation and human roles must complement each other.

The ethical dimensions of using AI in education remain contentious, especially regarding accountability. If gen-AI provides inaccurate or harmful content, determining responsibility becomes complex. This raises ethical concerns about accountability for errors or potentially harmful outcomes resulting from a reliance on AI-generated information [18]. Furthermore, integrating gen-AI into education represents a significant cultural shift, which some educators, institutions and parents may resist [19]. Concerns include the erosion of traditional education values, the challenge of keeping pace with rapid technological change and skepticism over AI's role in human-centric fields like education. Addressing these controversies involves balancing innovation with ethical considerations, equity and a commitment to ultimately fostering critical and independent thinkers.

1.3. Research Questions, Contribution and Structure

As gen-AI research in education expands, it is important to identify leading institutions that drive advancements, fund research and foster collaborations. Understanding institutional contributions may help in mapping the research ecosystem and recognizing centers of excellence. The first research question concerns this aspect:

RQ1: Which institutions have a major contribution to this field?

A major point of interest is the identification of current research topics in order to provide insights into prevalent discussions and areas of focus in the application of gen-AI in education. Uncovering underlying themes in a large corpus of research articles can be performed using the Latent Dirichlet Allocation (LDA) method. Therefore, the second research question is the following:

RQ2: What are the main topics extracted by LDA?

While some research topics receive significant attention, others may remain underexplored. Therefore, the third research question is as follows:

RQ3: Which topics can be considered insufficiently addressed?

The adoption of gen-AI in education, including ethical or academic policy decisions, is likely shaped by the public and scholarly sentiment toward this field. The fourth question aims to clarify this aspect:

RQ4: What is the sentiment on generative AI technologies in education?

The contribution of this research lies in providing a comprehensive bibliometric analysis of publications related to gen-AI and its applications in education. Specifically, it offers the following contributions (the corresponding research questions are given in parentheses):

- (RQ2, RQ3) The research organizes and synthesizes a growing body of knowledge, highlighting trends, emerging topics and best practices in the application of gen-AI in education.
- (RQ3) It pinpoints areas that are insufficiently studied, offering insights into where future research efforts should be directed.
- (RQ1) It identifies leading institutions, publications and researchers in the field, fostering opportunities for collaboration and the development of interdisciplinary research networks.
- (RQ2) It assesses the main latent topics of gen-AI applications in teaching and learning, contributing to the evolution of pedagogical approaches that have been validated through research.
- (RQ4) The research informs debates on ethical guidelines, addressing challenges such as privacy, intellectual property, bias and student data usage, thereby promoting responsible AI integration.

By answering these specific research questions, our research aims to offer actionable insights into the sentiment, trends and novelty of gen-AI in education.

This paper is organized into several sections. In this section, we provide an overview of the general context, challenges and motivation for analyzing education and AI, along with the objectives, controversies, research questions and our contribution to the state of the art. In the second section, a brief literature review on the use of gen-AI in education is provided. The proposed methodology is presented in Section 3. Section 4 is dedicated to the results and their interpretations, whereas in Section 5, the main conclusions are drawn.

2. A Review of Related Studies

This section synthesizes existing research on the application of gen-AI in education, highlighting trends and dominant themes, and identifies gaps to be addressed by our research. Its focus is on publications with similar aims—bibliometrics analyses and/or literature reviews on gen-AI/LLM's impact in the field of education.

The selected articles, published in 2023 and 2024, were identified by searching the Google Scholar database for the keywords “generative artificial intelligence”, “large language models”, “Chat*GPT”, “education”, “bibliometric”, and “review”. The timeframe was selected in order to reflect the research that followed the release of ChatGPT on 30 November 2022.

Due to differences in the scope and method of these studies, this section is structured into (1) reviews on AI in education, (2) bibliometric analyses on gen-AI and general education, (3) reviews (non-bibliometric) on gen-AI and general education and (4) reviews on gen-AI and specific education levels (K-12, higher education).

Looking at the broader scope of AI in education, Ivanova et al. [20] conducted a bibliometric analysis with a focus on teaching, revealing that ChatGPT and generative AI are prominent trends within AI in education and emphasizing their applications across various educational settings, highlighting both their supportive benefits and associated ethical challenges. Similarly, Qin et al. [21] identified ChatGPT as a key theme in AI in education (AIEd), noting its significant role in text-based applications while pointing out the underexplored potential of generative AI in multimedia domains such as images, videos and audio.

Moreover, gen-AI has rapidly emerged as a transformative technology in the educational domain [22], shaping diverse areas from pedagogy and assessment to academic integrity and institutional policies. Multiple bibliometric and systematic review studies, e.g., [23,24], underscore the dynamic growth of research publications focusing on ChatGPT and related LLMs in education.

Considering relevant bibliometric analyses related to the broader education domain (general education), we identified and analyzed 12 reviews. These are listed in Table 1 and described using a number of detailed attributes. From these articles, Scopus emerged as the most commonly utilized database, though several studies also drew from Web of Science, Google Scholar and Altmetric sources, sometimes in combination (e.g., WoS and Altmetric [25]). The reported search dates ranged broadly—from early 2023 to as recent as mid-2024—indicating ongoing and rapidly evolving research. Most authors adopted a bibliometric approach (often paired with systematic literature reviews or PRISMA-guided processes), while additional techniques included network analysis, data mining and analytic approaches or altmetrics to capture both academic and social media impacts.

In terms of observed research themes, a recurring observation across multiple reviews is the emergence of common thematic clusters. The interaction and communication with gen-AI-powered chatbots is explored in [26,27], emphasizing how conversational educational agents can enhance learning experiences and student engagement. The impact of gen-AI on teaching strategies, learner engagement and pedagogy is regarded as central [26]. Authors highlight opportunities for personalized instruction and the transformation of

course-content delivery. In connection with assessment and academic integrity, concerns about plagiarism, cheating and the reliability of AI-assisted assignments are frequently discussed [27–29]. Researchers advocate for novel assessment designs to address these challenges [28]. Various reviews note the ethical issues, privacy concerns and potential biases in AI [26,30]. Researchers call for AI literacy to empower learners and educators to use ChatGPT responsibly.

In terms of general vs. specific education domains, the findings of Chang et al. [31] indicate that the majority of research in this area focuses on general applications of ChatGPT, with substantial studies also targeting English learning, STEAM education and medical education. ChatGPT's role in medical and nursing education features prominently, driven by the model's potential in patient simulation, diagnostic training and content generation [32,33]. Additionally, computing education emerges as a field that is especially sensitive to the dual role of ChatGPT as both a powerful tool for code generation and a potential means of academic misconduct [28].

With regard to geographic distributions, multiple bibliometric analyses identify the United States, China, Indonesia and Australia as leading contributors [24,32,34]. Research disciplines range broadly—spanning computer science, engineering, medical education and social sciences—underscoring the interdisciplinary nature of ChatGPT-related studies [23,25].

Besides the above-mentioned bibliometric reviews, a number of early reviews have followed scoping, systematic or narrative approaches in order to identify the practical applications and ethical implications of LLMs like ChatGPT in educational settings. Yan et al. [12] identified 53 use cases across nine categories, including grading, teaching support and personalized learning, highlighting the versatility of LLMs in automating educational tasks while emphasizing the need for updated models and human-centered designs. Grassini [35] and Ipek et al. [36] explored the benefits of ChatGPT in enhancing personalized learning and facilitating complex academic tasks, alongside concerns about accuracy, academic dishonesty and data privacy. These reviews advocate leveraging generative AI's strengths while addressing ethical issues through robust guidelines and innovative strategies to ensure responsible and effective integration in educational settings.

Still other review papers have focused on specific levels of education (K-12 or higher education). Zhang and Tur [37] focused on K-12 education, demonstrating that ChatGPT empowers educators through curriculum development and personalized student support but also poses challenges to academic integrity and output quality. Ansari et al. [38] conducted a systematic review of 69 studies on ChatGPT in higher education, revealing its widespread use in supporting teaching and learning activities while highlighting concerns related to accuracy, academic integrity and student development. They propose a comprehensive framework for universities to effectively integrate ChatGPT. Similarly, Peláez-Sánchez et al. [39] reviewed 83 articles on LLMs within the framework of Education 4.0, emphasizing their potential to create more autonomous and interactive learning environments. However, they stress the importance of human oversight and addressing ethical and legal challenges to ensure that AI integration aligns with educational objectives. Dempere et al. [40] emphasize the transformative potential of ChatGPT in higher education, including research support and automated grading, alongside risks such as privacy breaches and misinformation. Imran and Almusharraf specifically examined ChatGPT as a writing assistant, highlighting its role in facilitating academic writing while addressing the necessity for updated policies to maintain academic integrity.

Table 1. Papers with bibliometric analyses of gen-AI in education.

Reference	Publication Date/Early Access Date	No. of Publications Analyzed	Type of Analysis	Databases Used	Date of Search (Year.Month)	Search Terms/Keywords
[34]	November 2024/July 2024	45	Bibliometric analysis, systematic literature review	N.A.	2024.05	N.A.
[23]	September 2023	207	Bibliometric analysis, content analysis	Scopus	2023.07	"education", "GAI", "generative artificial intelligence", "GPT-4", "ChatGPT", "AlphaCode", "GitHub Copilot", "Bard"
[27]	November 2024/May 2024	47	Bibliometric analysis	Scopus	2023.05	"ChatGPT and teaching learning in universities" OR "effect of ChatGPT in higher education institution" OR "ChatGPT and student assessment in higher education" OR "ChatGPT and academic integrity" OR "ChatGPT and teaching pedagogy in higher education institution" OR "ChatGPT and cheating student course assignment" OR "ChatGPT and teaching in higher education" OR "implications of ChatGPT in higher education institutions" OR "ChatGPT and evaluation criteria in higher education institution" OR "ChatGPT in universities" OR "ChatGPT and student learnings"
[26]	2023	513	Data mining: t-SNE analysis, text mining and social network analysis	Scopus	N.A.	Article title: "generative AI" OR "generative artificial intelligence" OR "gen AI" OR "GPT" OR "chatbot" OR "conversational agent" OR "chatgpt" OR "bard" OR "large language model" AND Title, abstract, keywords: "education" OR "school" OR "colleague" OR "Tertiary Education" OR "teacher" OR "student"
[30]	2024	109	Bibliometric analysis	Scopus	N.A.	N.A.
[33]	2024	212	Bibliometric analysis	Scopus	2023.07	"chatgpt" AND ("education" OR "instruction" OR "teaching" OR "learning")
[22]	December 2023	74	Bibliometric analysis, systematic literature review, network analysis	Google Scholar	N.A.	N.A.
[28]	August 2024	64	Bibliometric analysis, systematic literature review	Scopus	2024.01	"generative artificial intelligence" or "generative AI" or "Gen AI", AND; "higher education" or "university" or "college" or "post-secondary", AND; "impact" or "effect" or "influence"
[29]	2023	51	Bibliometric analysis	Web of Science	2023.06	(Title = ("ChatGPT")) AND (Subject Area = ("educational research"))

Table 1. Cont.

Publication Reference	Publication Date/Early Access Date	No. of Publications Analyzed	Type of Analysis	Databases Used	Date of Search (Year.Month)	Search Terms/Keywords
[25]	August 2024/ February 2024	175	Combines altmetrics and bibliometrics	WoS, Altmetric	2023.08	"ChatGPT" + subjects: "education educational research", "education scientific disciplines", and "psychology education". Altmetric data: the subjects were limited to "education", "education policy, sociology and philosophy", "education systems", "specialist studies in education", and "engineering practice and education".
[24]	2024	58	Bibliometric analysis	Scopus	2024.03	"ChatGPT" AND "education" OR "artificial learning" OR "education technology"
[32]	January 2024	82	Bibliometric analysis	Scopus	N.A.	"ChatGPT" AND "education"

Considering the attributes of previous studies, we observe a series of limitations related to not accounting for newer publications and variations in the scope of analysis due to non-standard search terms or keywords used to retrieve papers—an aspect also highlighted by Wong et al. [25]. Specifically, 6 of the 12 analyzed reviews used only "ChatGPT" as a search term in reference to LLMs/gen-AI, although this is just one specific LLM-based application and is not synonymous with LLM technology in general.

The current bibliometric research addresses critical gaps in the existing literature in the following ways: (1) By capturing the most recent developments in gen-AI in education, encompassing publications up to the moment of elaboration. (2) Unlike earlier reviews that often focus narrowly on ChatGPT alone, our investigation includes papers addressing the broader field of generative AI and LLMs. Furthermore, (3) by incorporating sentiment analysis and topic modeling, we move beyond merely quantifying research trends to examining the perceptual dimensions of AI adoption. These aspects position our research to advance the current discourse and to guide future research into this rapidly evolving field.

3. Methodology

The employed methodology follows a comprehensive workflow for analyzing bibliometric and content data retrieved from Web of Science (WoS), as shown in Figure 1. It begins with the data collection process, which aims to collect data about publications relevant to the field of interest—generative AI in education.

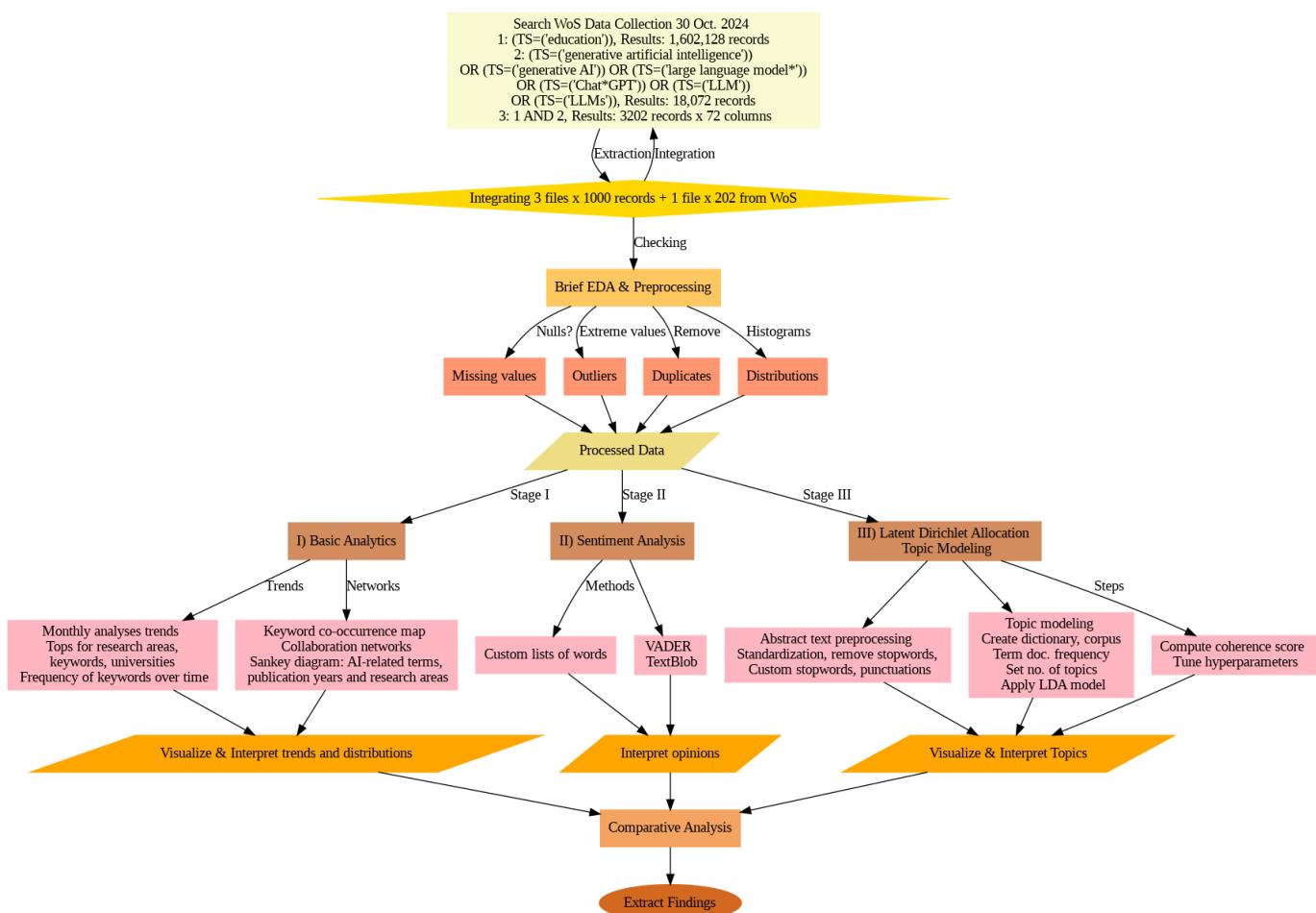


Figure 1. Flowchart diagram.

As the intention was to include all relevant and available WoS data without undue bias, the initial search set as few limitations as possible; therefore, there was no restriction based on the publication timeframe, type of publication, research area, etc. The search terms were applied only with respect to the publication topic. We used the WoS “Topic” (TS) field, which included the title, abstract and author keywords in the search. The issue of possibly irrelevant results was dealt with in the exploratory data analysis and preprocessing phase, as will be described later.

On 30 October 2024, three searches were conducted in WoS, as shown in Table 2. The first search, TS = (“education”), focused on identifying records related to education by looking for the term “education” in fields such as titles, abstracts and keywords. This search returned 1,602,128 records, highlighting the vast amount of research available on educational topics across various disciplines. The second search, TS = ((“generative artificial intelligence”) OR (“generative AI”) OR (“large language model*”) OR (“Chat*GPT”) OR (“LLM”) OR (“LLMs”)), targeted publications discussing gen-AI. The use of multiple terms and the logical operator OR ensured comprehensive coverage of this topic. This query returned 18,072 records, reflecting a more focused but still substantial body of research on generative AI technologies. The third search, combining the results of the first two with 1 AND 2, focused on the intersection of education and generative AI. This retrieved records that mentioned both topics, ensuring relevance to their combined application. Examples of such research include studies on AI-driven tools for teaching, the impact of ChatGPT on educational practices or ethical considerations in using AI in learning

environments. These searches illustrate a strategy for exploring emerging intersections between broad topics (education) and specific innovations (generative AI).

Table 2. Set of searches in the WoS Core Collection database.

Search	Keywords Search Date: Wed 30 October 2024	Results (Records)
1	(TS = (education))	1,602,128
2	(TS = ("generative artificial intelligence")) OR (TS = ("generative AI")) OR (TS = ("large language model*")) OR (TS = ("Chat*GPT")) OR (TS = ("LLM")) OR (TS = ("LLMs"))	18,072
3	1 AND 2	3850

Table 3 provides a detailed overview of the metadata associated with the research articles, including their titles, sources, keywords, abstracts, author affiliations, citation metrics, usage statistics, publication details and classifications within the WoS database.

Table 3. Data column descriptions.

Column Name	Description
Article Title	The title of the research articles, highlighting their primary focus.
Source Title	The journal or publication source where the article was published.
Author Keywords	Keywords provided by the authors, summarizing the main topics of the research.
Keywords Plus	Additional keywords derived from indexing services.
Abstract	A summary of the research article, including objectives, methods and findings.
Affiliations	Institutions or organizations the authors were associated with.
Cited Reference Count	Total number of references cited in the article.
Times Cited, WoS Core	Number of citations in the Web of Science Core Collection.
Times Cited, All Databases	Total number of citations across all databases.
180 Day Usage Count	Number of accesses to the article in the last 180 days.
Since 2013 Usage Count	Total number of accesses to the article since 2013.
Publisher	The organization responsible for publishing the article.
Publication Date	The specific date or month the article was published.
Publication Year	The year the article was published.
WoS Categories	Categories under which the article was classified in the Web of Science.
Web of Science Index	Indexing category for the article in the Web of Science.
Research Areas	Broader areas of research the article pertained to.

The next phase involved data preprocessing, which started with Exploratory Data Analysis (EDA) to clean and structure the data. This included identifying and addressing missing values, removing duplicates and analyzing distributions using histograms. The goal was to produce a processed dataset that is ready for detailed analysis.

The inspection of the dataset showed that the database search returned some records which did not conform to the intended scope of the study (e.g., the acronym LLM being used with different meanings or publications on gen-AI that had a very weak or incidental focus on education). Therefore, a screening process was further applied, resulting in a dataset of 3506 publications. The screening consisted of excluding the following records:

- i. All publications before 2019, which, after manual verification, were found not to be relevant to the study (28 records);
- ii. Publications that contained “LLM” but no other specific gen-AI keywords—27 records;
- iii. Publications with a WoS Category (research field) not including education AND having no (zero) education-related keywords in the title and author keywords AND having less than two education-related keywords in the abstract—280 records;
- iv. Duplicate publications in terms of author(s) and title—10 records.

The main analysis contains the application of bibliometric analysis methods for performance analysis, science mapping, citation analysis, clustering and visualization (as described by Donthu et al. [10]), sentiment analysis and topic modeling and is divided into three stages. Stage I focuses on basic analytics, such as identifying publication trends over time, analyzing keywords and visualizing collaboration networks. Tools like Sankey diagrams are used to map relationships between AI-related terms, research areas and publication years. This stage provides an understanding of trends and distributions in the dataset. Stage II conducts sentiment analysis using tools like VADER and TextBlob to interpret the opinions expressed in the abstracts. Custom word lists are also utilized to refine the sentiment analysis, helping to understand the attitudes related to the studied topics. Stage III applies Latent Dirichlet Allocation (LDA) for topic modeling. This involves preprocessing text data, creating dictionaries and corpora, and performing topic modeling to uncover thematic structures. Hyperparameter tuning, including coherence score optimization, ensures the best results.

The following software tools and libraries were used for data processing and analysis: Python 3.11, pandas 2.2.2, NLTK 3.9.1, TextBlob 0.15.3 and Gensim v.4.3.3.

The final steps synthesize the findings from all stages. Results are visualized and interpreted to provide a comprehensive understanding of the trends, opinions and thematic insights. Comparative analyses across these dimensions further refine the conclusions. Ultimately, the workflow extracts actionable findings, shedding light on the intersection of education and generative AI, offering valuable insights for research and practice.

4. Results and Discussion

This section presents a comprehensive analysis of the collected data, structured into three sub-sections: “Trends and Co-occurrences in Data”, “Sentiment Analysis” and “Topic Modeling with LDA”. “Trends and Co-occurrences in Data” explores the time dynamics of publications, their keyword patterns, evolution and co-occurrence, and the most influential institutions that drive academic research within the field of generative AI in education. “Sentiment Analysis” examines the prevailing attitudes expressed in the literature, providing insights into the overall perception of ChatGPT and other generative AI technologies as reflected by the research community. Finally, “Topic Modeling with LDA” identifies and categorizes the underlying themes and subjects across the reviewed studies.

4.1. Trends and Co-Occurrences in Data

In the timeframe between 2019 and 2024, a total number of 3506 publications were identified (Figure 2); when looking at the time evolution, it is obvious that there was an explosion in the number of publications in 2023 and 2024, which was expected, considering the ChatGPT public launch at the end of November 2022.

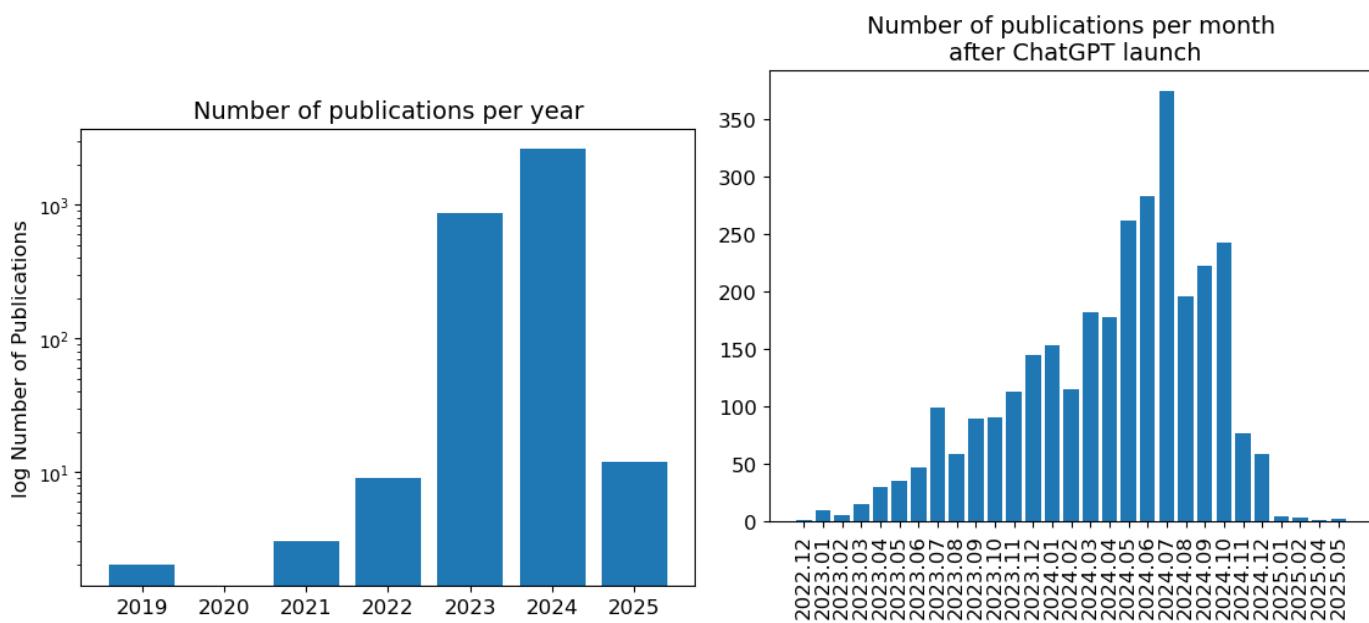


Figure 2. Number of publications by year (2019–2025) and by month (December 2022–May 2025).

For a more fine-grained analysis, we looked at the publication month, which was available by gathering data from the columns “Publication Date” (for articles published in journals), “Conference Date” (for conference papers); if these were not available, the “Early Access Date” was used. No month data were available for 10.78% of the observations. Given that for the years 2019–2022, the monthly publication count was very low (two or less; zero for most months), this analysis focused on the timeframe starting from December 2022, which is the period directly following ChatGPT’s launch. The results in Figure 2 (right side) show almost continuous growth from month to month, with a peak of almost 400 publications in July 2024. The decrease in the months following July 2024 is likely due to the fact that not all the relevant publications had been indexed in the Clarivate WoS database at the time of the data collection in mid-November 2024. The data that existed for “future” months (November 2024–May 2025) were related to a number of early-access publications. It can also be observed that, since the 2024 data does not cover the entire year (up to the end of October), the actual number of publications at the end of 2024 is expected to be even higher than the one currently recorded.

With regard to the Clarivate WoS research areas, “Education & Educational Research” and “Computer Science” were far ahead of other areas, with 1576 and 910 publications, respectively. Other areas showing an interest in the topic of concern were “Engineering”, “General and Internal Medicine”, “Surgery” and “Health Care Sciences & Services”, with the latter three covering the broader area of health sciences (as in Figure 3). It has to be noted that one publication may belong to more than one research area and subject category, so there is some overlap in the publication counts. The publications in the areas of education, computer science and medicine were further grouped by subject category, providing a somewhat more detailed view of the corresponding subjects. There were two categories for education, five categories for computer science and three health-related categories in the top ten.

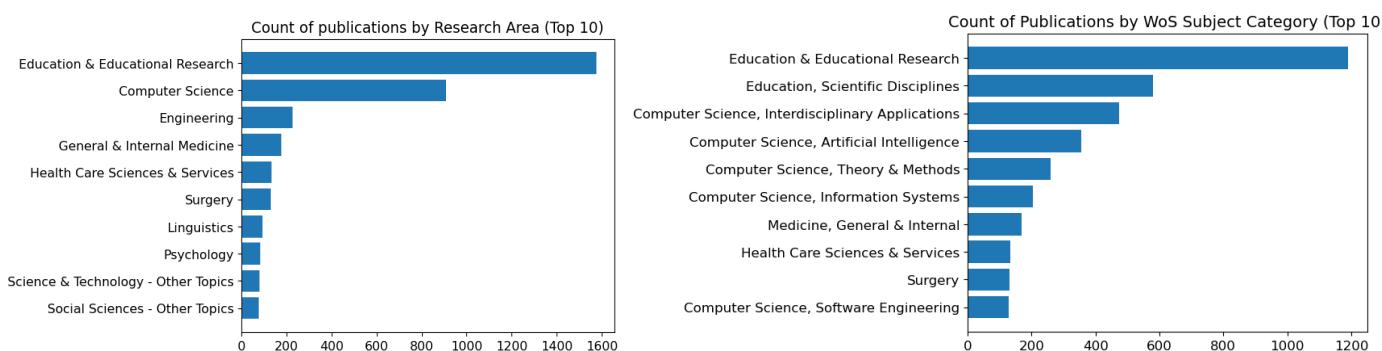


Figure 3. Top ten research areas and subject categories with the most publications.

The list of top keywords by their count shows that a technology focus predominates, with “ChatGPT” and “artificial intelligence” and similar terms ranking the highest, as shown in Figure 4. Education, higher education and medical education are mentioned as keywords in 200 to almost 400 publications. Each publication may be associated with multiple keywords.

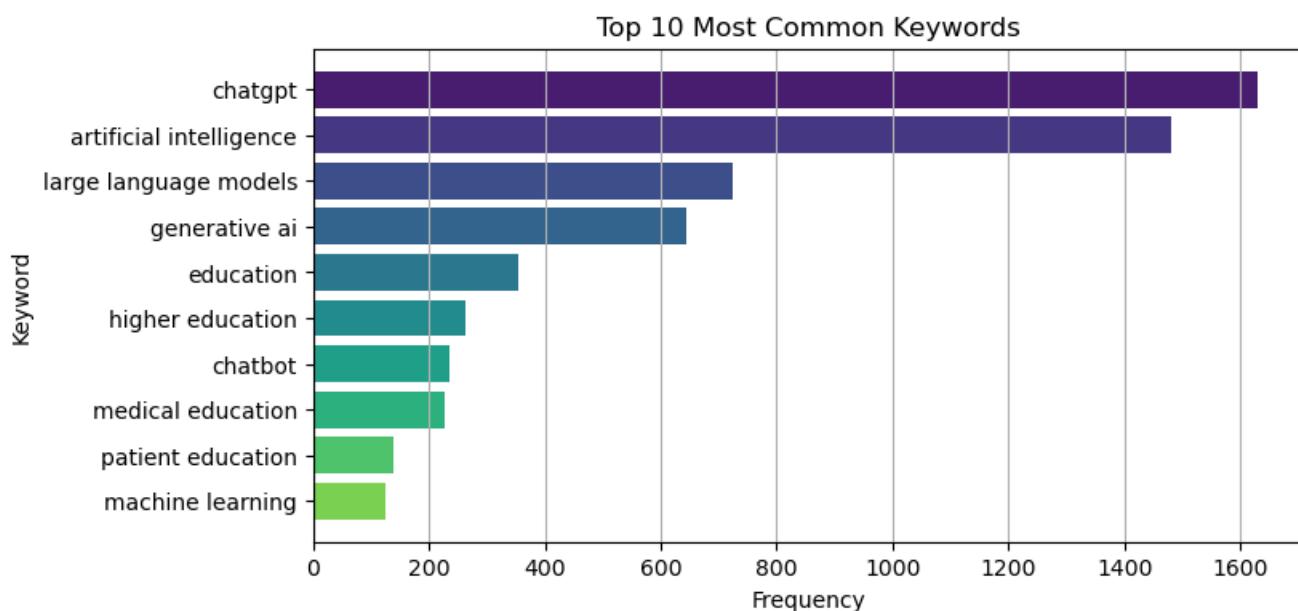


Figure 4. Top ten most frequent keywords.

Looking at the keyword count over time (Figure 5), we observe that the terms “ChatGPT” and “artificial intelligence” had the highest growth rate during 2023–2024, followed by “large language models”, “generative AI” and “education”. The timeframe was chosen to focus on the months with a consistent publication count (see also Figure 2), so that the keyword data could be summarized. As already mentioned, the counts after July 2024 may not reflect the actual number of publications due to the WoS indexing lag.

The keyword co-occurrence map of the top ten keywords (Figure 6), shows the number of publications containing a given combination of two keywords. “Education”, “higher education” and “medical education” occurred together the most with the top technology terms “ChatGPT” and “generative AI”.

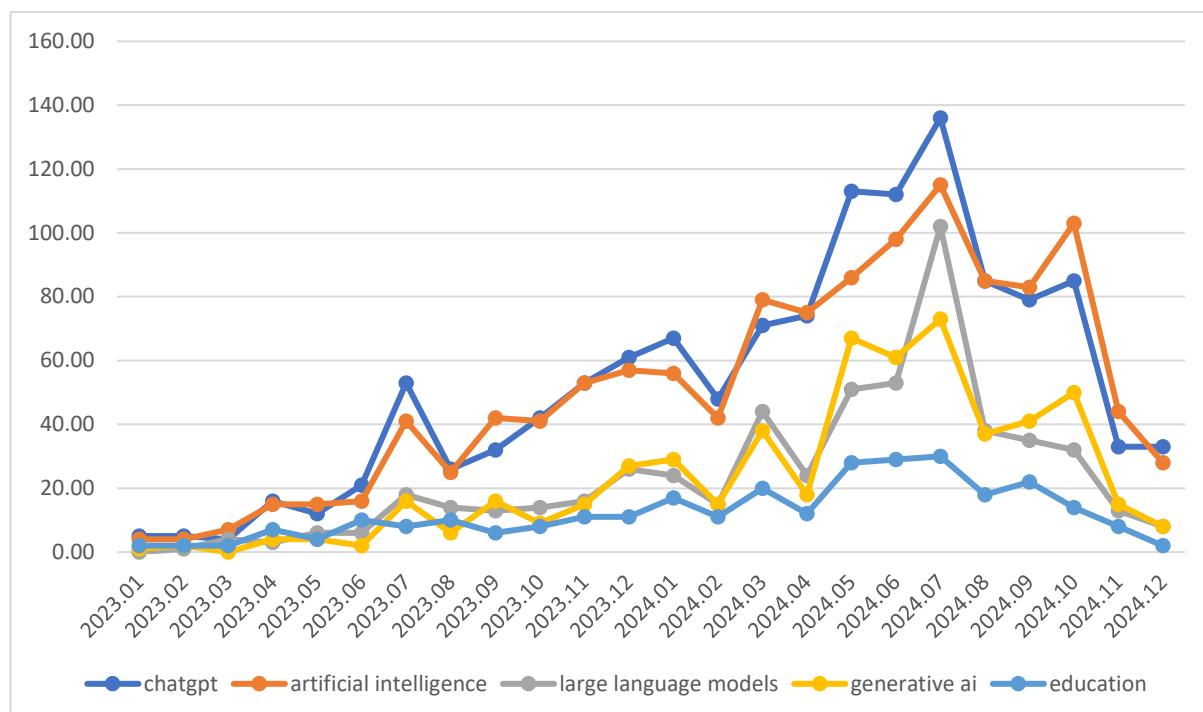


Figure 5. Keyword frequency over time for the period January 2023–December 2024.

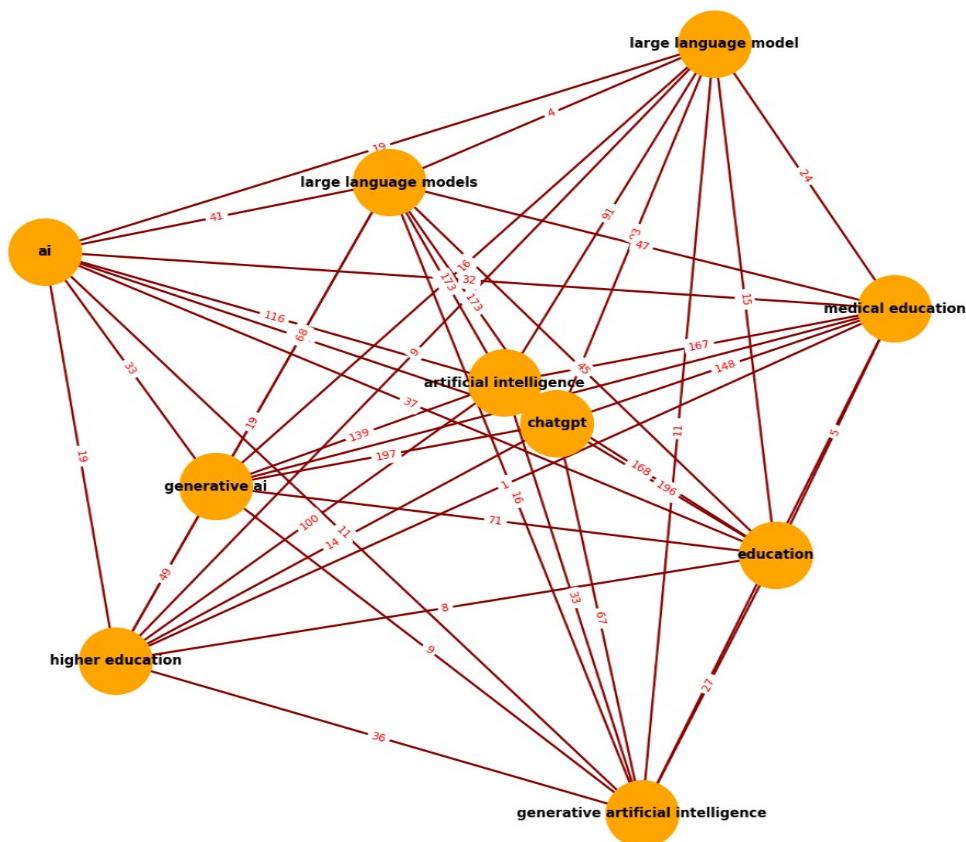


Figure 6. Keyword co-occurrence map.

The Sankey diagram in Figure 7 was created in order to visualize the mapping between AI-related keywords, research areas and publication years. The link width is proportional to the number of publications that feature the corresponding attribute values. In terms of AI keywords related to publication years, given the fact that in 2024, the number of

publications of interest had more than tripled compared with 2023, the keyword “generative AI” saw an even higher rate of growth, as most publications with this keyword were linked to 2024. The diagram begins on the left with broad topics such as “artificial intelligence” and “generative artificial intelligence” and shows how these topics diverged and evolved over time. By 2024, AI applications had branched into specialized fields, including education and educational research, computer science, engineering, psychology, medical informatics, health care sciences, general medicine and surgery. With regard to the top research areas, the number of publications rose sharply in 2023 compared with the previous years for most areas.

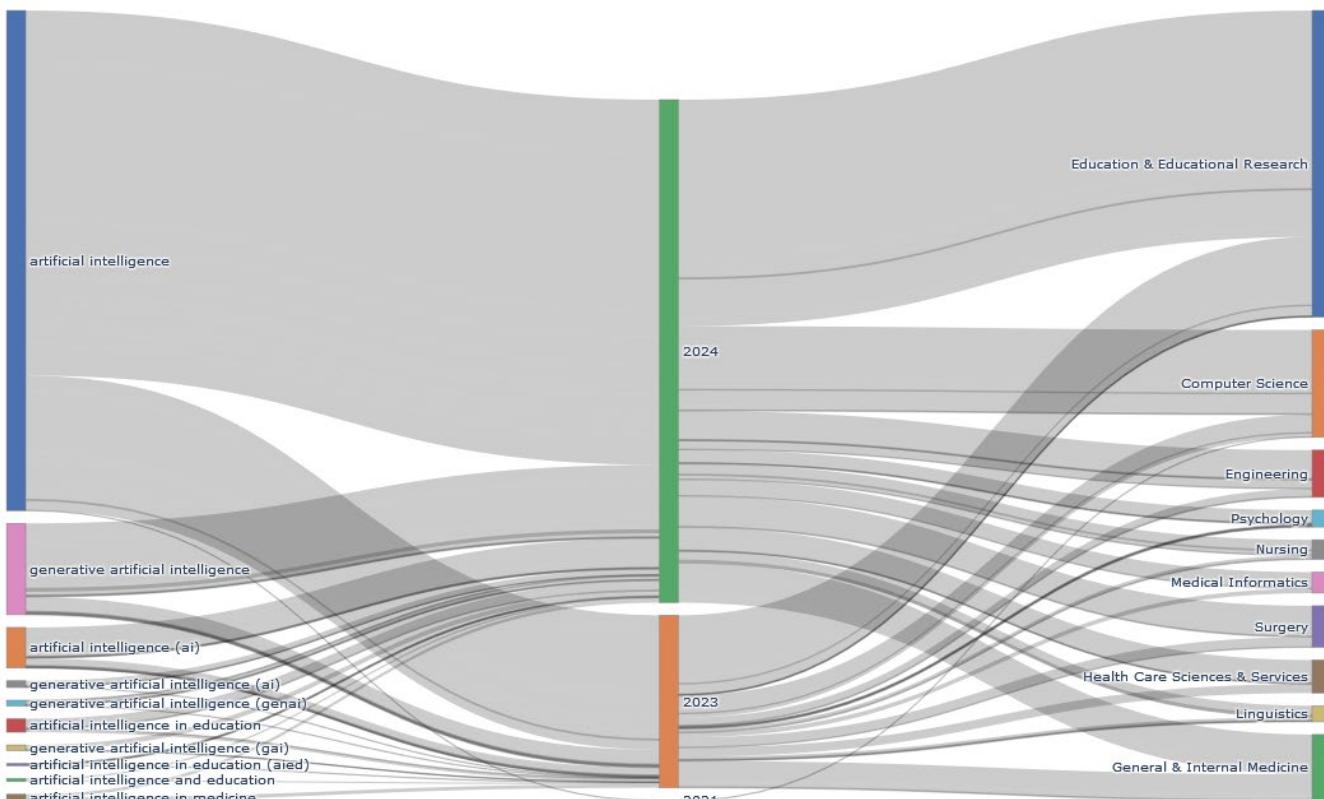


Figure 7. Sankey diagram mapping keywords, publication year and research area.

The ranking of the higher education institutions to which publication authors were affiliated was calculated based on the publication count. The University of California system had the highest number of publications, 78, representing 2.22% of all the observations. The top 10 institutional affiliations (Figure 8) shows that most of the leading universities were located in the United States, with exceptions such as the University of London (3rd place) and the University of Toronto (10th place), which was closely followed by the Singapore and Hong Kong Universities. US university systems usually include a number of campuses, as well as research laboratories and other facilities and accommodate a large number of researchers (e.g., ten campuses in the University of California system, with 265,000 faculty and staff, according to its website).

Figure 9 shows the collaboration connections between these ten institutions (represented as nodes) in relation to the co-authorship of publications. The number of jointly authored publications is shown on the connecting arc. The strongest collaborations can be observed between the Universities of Harvard and Stanford (34 joint publications), Harvard and Toronto (32) and Harvard and the University of California system (21).

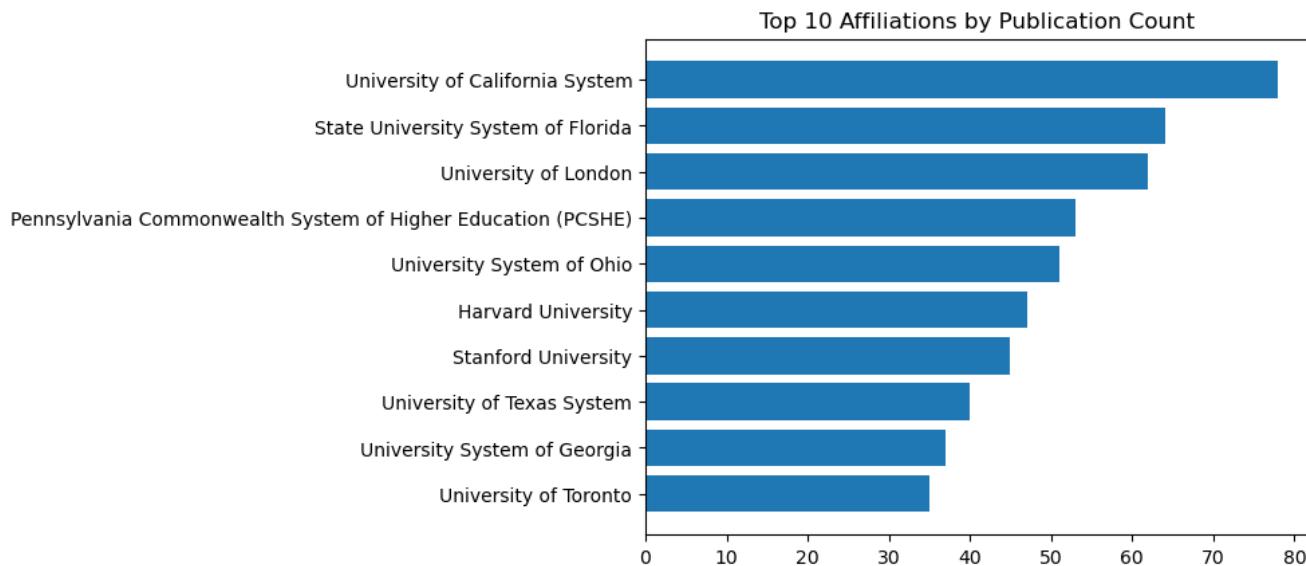


Figure 8. Top ten affiliations by publication count.

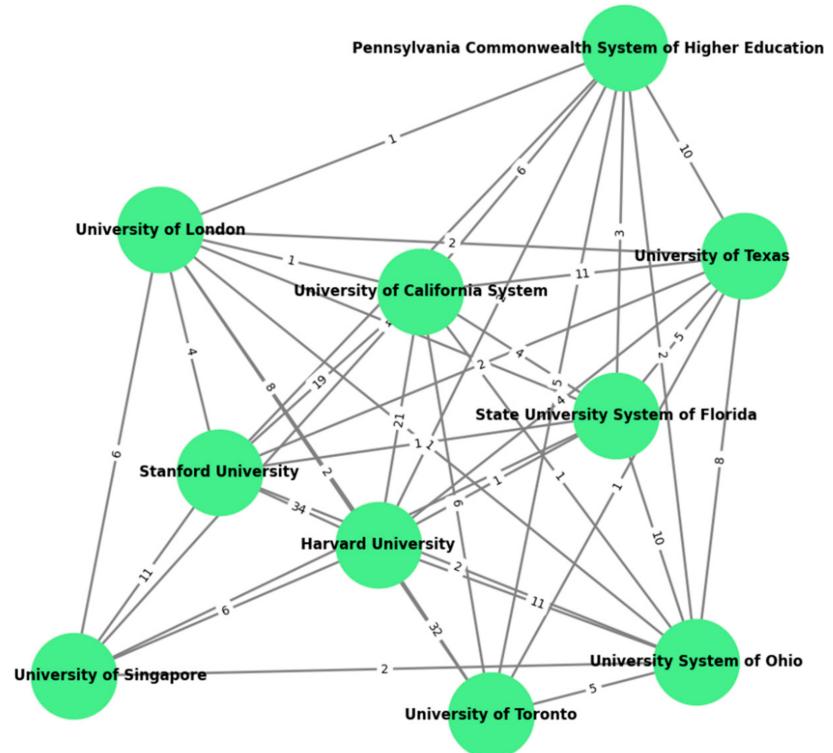


Figure 9. Collaboration network for the top ten institutions.

The ranking by number of citations shows that a few publications were highly cited, ranging from 926 times for the first position, down to 254 for the tenth position (see Figure 10). Considering the time required to publish research studies, the list is unsurprisingly biased toward early publications, as there was enough time for them to also be referenced in later publications. Seven of the top ten publications appeared no later than April 2023, and the others were published from June through to August 2023. The top article in terms of citations, although published in August 2023, adopted a multidisciplinary approach that covered, besides education, a broad range of areas that are impacted by gen-AI; hence, it had greater potential for being cited by publications in other research areas.

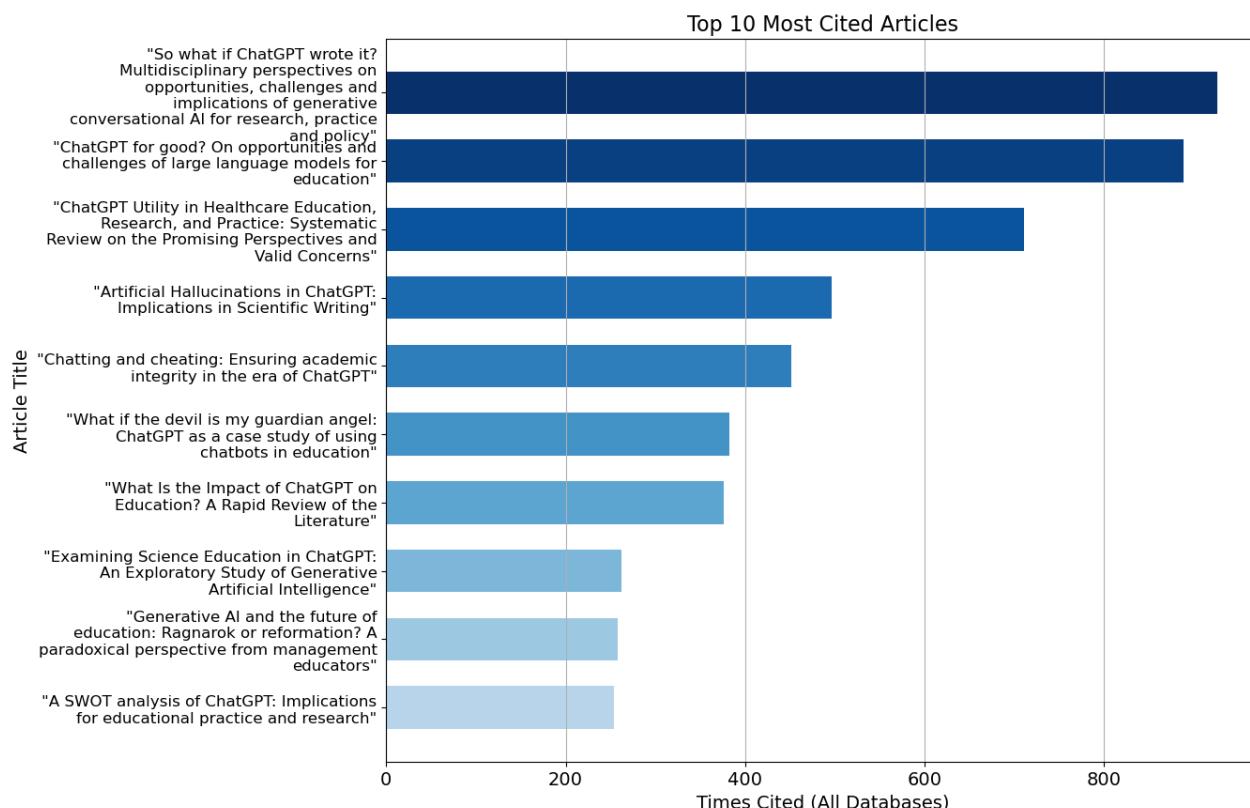


Figure 10. Top ten most cited publications.

4.2. Sentiment Analysis

Sentiment analysis (SA) techniques are generally used to determine positive/negative/neutral opinions or attitudes in a text. In this case, a number of sentiment analysis methods were applied to the publication abstracts in the dataset in order to measure the overall attitude of optimism/skepticism with regard to gen-AI in education, as well as possible consensus/controversy in this field. We are also interested in investigating the applicability of such techniques to scientific literature, considering that it is usually written in an objective and neutral manner, but the negative or positive trends over time have yet to be identified.

The timeframe for this analysis, January 2023–November 2024, includes months with a consistent number of publications (see also Figure 2), for which there are sufficient data available to apply SA methods. Within this timeframe, the selection of publications having available abstracts resulted in a subset of 2863 publications.

The first SA technique employed is a custom algorithm based on counting positive and negative words (from predefined lists) that may appear in each abstract. The SA indicators are computed for all publications in the dataset, as well as for subsets of publications by the time of publications, with the aim of identifying a possible change in sentiment over time.

The results (in Figure 11) show that while an abstract contains approximately 200 words, most of these words are rated as neutral; positive words are found in about four out of five (~80%) publications, while negative words are less frequent, being found only in about one out of five (~20%) cases. We may conclude, therefore, that while a neutral writing style predominates, there are more publications that exhibit a positive opinion/attitude than those expressing a negative one. The SA based on the time of publication shows a significant increase in positive sentiment from 0.705 in 2023 to 0.816 in 2024 (as shown in Table 4).

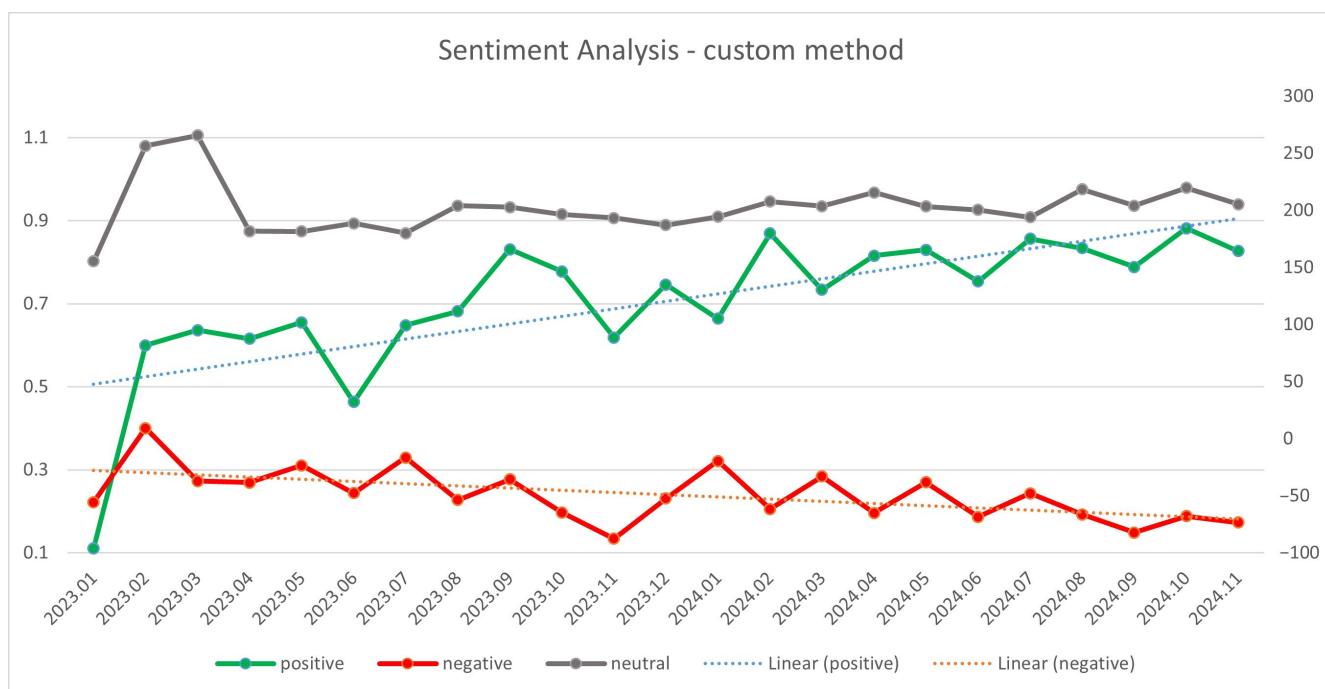


Figure 11. Sentiment in publication abstracts over time using the custom method.

Table 4. Sentiment analysis results by a custom list of words.

Word Count (Means)			
	Positive	Negative	Neutral
2023	0.705	0.237	199.080
2024	0.816	0.218	208.328

The monthly SA chart confirms a trend of increase in positive sentiment, while the negative sentiment shows a slight decrease over time.

Two other SA methods are used as alternatives for evaluating opinions and attitudes: VADER and TextBlob. VADER, as part of NLTK 3.9.1, is a lexicon- and rule-based SA tool that evaluates the probabilities for a text to express a negative, positive or neutral sentiment [41]. Its main indicator, “compound”, measures the overall sentiment, which is considered positive for values greater than 0.05. The model was trained specifically for evaluating social media posts. TextBlob version 0.15.3 is a Python library for NLP, which includes an SA feature [42]. TextBlob computes the “polarity” and “subjectivity” of a given text; the polarity indicator ranges from -1 (negative sentiment) to $+1$ (positive sentiment), while the subjectivity ranges between 0 (very objective) and 1 (very subjective/personal opinion). The VADER results confirmed the initial findings, showing that a neutral sentiment predominated (0.86, neutral score; 0.77, compound), while positive attitudes scored higher than negative ones (0.11 positive vs. 0.025 negative). The evolution over time in Table 5 and Figure 12 is also similar, showing a slight increase in positive sentiment in 2024 vs. 2023.

Table 5. SA results by VADER.

VADER Metrics (Means)				
Year	Neg	Neu	Pos	Compound
2023	0.026	0.865	0.109	0.775
2024	0.024	0.863	0.113	0.813

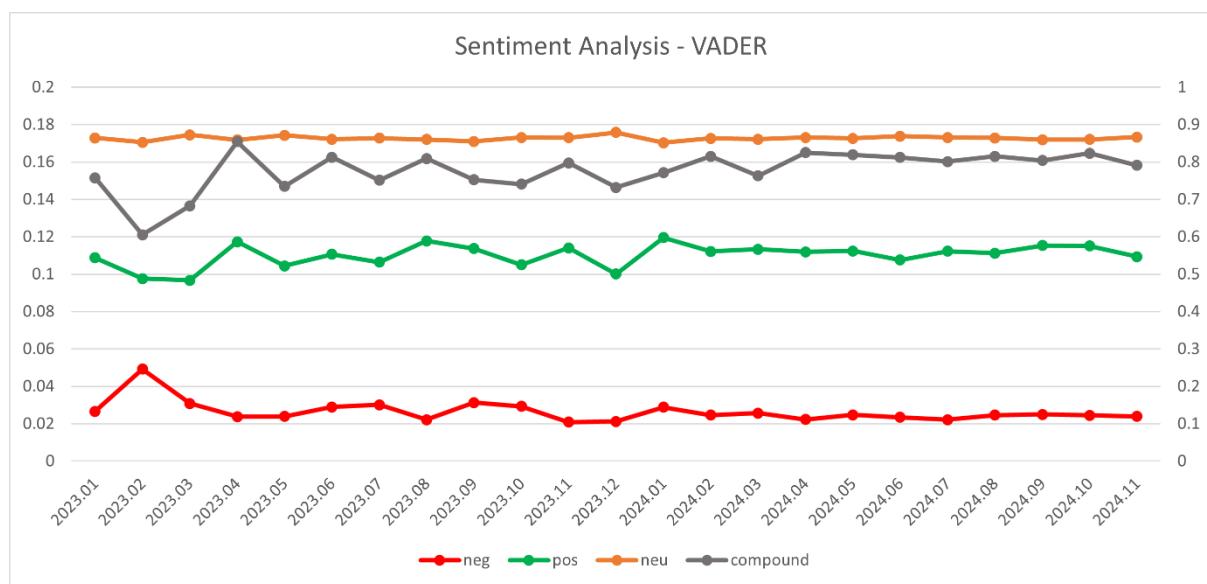


Figure 12. Sentiment in publication abstracts over time using the VADER method.

The TextBlob results show that the polarity registered positive values of approx. 0.09, indicating a slightly positive sentiment, while subjectivity had a value of 0.47, which can be interpreted as the abstracts being neither strictly objective nor very subjective. The time-related analysis was consistent with the other methods in measuring a higher polarity (positive sentiment) for publications in 2024 than those in 2023, while the subjectivity level showed a minor decrease, as shown in Table 6 and Figure 13.

Table 6. SA results by TextBlob.

TextBlob Metrics (Means)		
Year	Polarity	Subjectivity
2023	0.086	0.477
2024	0.092	0.469

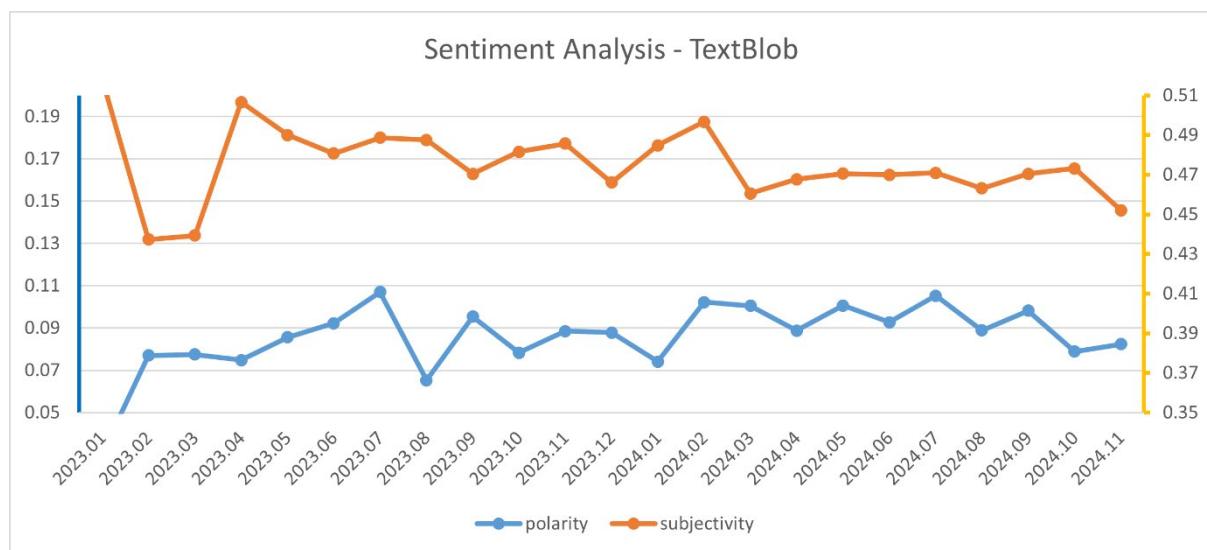


Figure 13. Sentiment in publication abstracts over time using the TextBlob method.

4.3. Topic Modeling with LDA

The LDA algorithm was applied to the publication abstracts as an unsupervised technique that clusters a corpus of documents into topics based on their similar word distributions. This enables researchers to discover common latent topics for sets of documents by using domain knowledge and judgement in order to assess the interpretability of the identified topics.

In order to apply the LDA algorithm, implementation within the Gensim library was used [43]. Out of the publications' dataset, 3256 publications that contained an abstract were selected. The data were preprocessed in order to convert all words to lowercase, remove non-alphabetic tokens and stopwords and to bring words to their common root by lemmatization. Then, the LDA algorithm was applied, setting the alpha and beta priors as symmetric and equal to $1.0/\text{num_topics}$, where num_topics is the chosen number of topics for clustering documents.

The C_v topic coherence measure [44], also implemented in the Gensim package, was used as a metric for comparing LDA models that differed in terms of the num_topics hyperparameter. Experiments with num_topics ranging from 2 to 20 led to the choice of $\text{num_topics} = 5$; the resulting model thus reaches a coherence score of 0.49 (values range between 0 and 1, with higher values indicating a higher coherence).

The LDA results were analyzed using a dedicated visualization tool: pyLDAvis (Python LDA Visualization [45]). Figure 14 contains topic circles on the left side, with the circle areas being proportional to the weight of the topic across the entire corpus. The topic circles are placed in a two-dimensional coordinate system based on the multidimensional scaling of the inter-topic distance matrix. The horizontal bar chart on the right side shows the overall term frequencies as blue bars and the term frequencies within a selected topic as red bars; the terms are sorted in descending order, by their relevance for topic 1, as defined in [45].

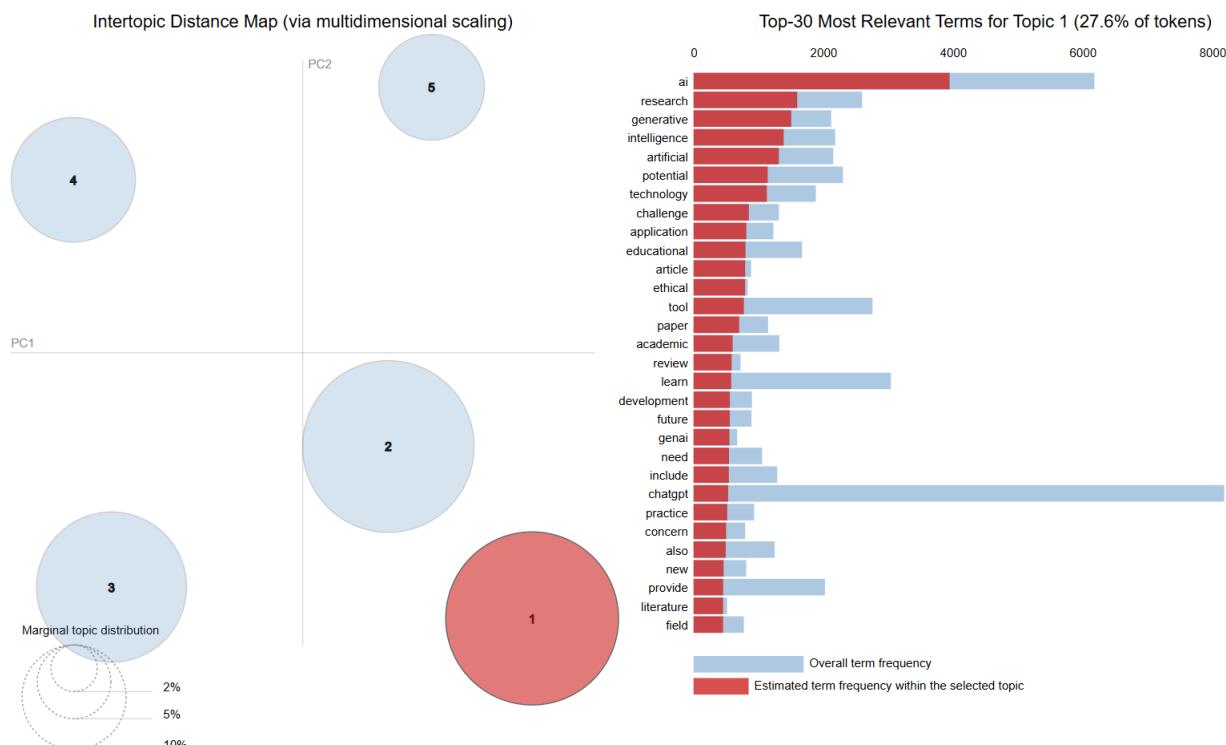


Figure 14. LDA visualization for Topic 1.

Figure 14 shows that five topics can be clearly differentiated—the circles are large in size and are placed far apart on the inter-topic distance map. In terms of the topic interpretability, the topic content is inferred from the most relevant terms of each topic, as shown in Table 7.

Table 7. Topic modeling results from LDA.

Topic No.	Topic Weight	Top 10 Terms and Weights
1	27.6%	0.038 * "ai" + 0.015 * "research" + 0.015 * "generative" + 0.013 * "intelligence" + 0.013 * "artificial" + 0.011 * "potential" + 0.011 * "technology" + 0.008 * "challenge" + 0.008 * "application" + 0.008 * "educational"
2	27.1%	0.041 * "student" + 0.041 * "chatgpt" + 0.023 * "study" + 0.016 * "ai" + 0.016 * "learn" + 0.012 * "tool" + 0.010 * "teacher" + 0.008 * "write" + 0.008 * "research" + 0.008 * "high"
3	20.8%	0.025 * "model" + 0.020 * "language" + 0.016 * "student" + 0.015 * "llm" + 0.012 * "large" + 0.011 * "learn" + 0.010 * "generate" + 0.008 * "task" + 0.008 * "feedback" + 0.007 * "prompt"
4	14.3%	0.034 * "chatgpt" + 0.024 * "patient" + 0.023 * "response" + 0.015 * "information" + 0.015 * "score" + 0.012 * "level" + 0.009 * "question" + 0.009 * "provide" + 0.009 * "study" + 0.009 * "health"
5	10.2%	0.052 * "question" + 0.045 * "medical" + 0.042 * "chatgpt" + 0.022 * "answer" + 0.019 * "performance" + 0.016 * "clinical" + 0.015 * "model" + 0.014 * "study" + 0.012 * "examination" + 0.011 * "accuracy"

The first topic, Gen-AI's impact in education and research, with a weight of 27.6%, revolves around the integration of AI, especially ChatGPT, into education and research, highlighting both its potential and its challenges, risks and limitations as well as ethical challenges.

The second topic, ChatGPT as a tool for university students and teachers (27.1%), explores how AI tools are being utilized by students and educators to enhance learning and to conduct research. The focus is on the supportive role of these tools in higher education and their broader impact on understanding AI itself.

The third topic, Large Language Models and prompting in computing education, carrying a weight of 20.8%, focuses on the use of LLMs in programming and code generation for educational purposes. It highlights their applications in creating programs, responding to prompts and their utility in educational contexts, particularly for students. Key terms such as "generate", "prompt", "task" and "feedback" suggest an emphasis on the interactive capabilities of LLMs for creative and computational tasks.

While both topics 4 and 5 are related to ChatGPT's role in medical/healthcare education and research, the former focuses on applications of ChatGPT in patient education (14.3%), especially its ability to provide responses to patient inquiries, while the latter is concerned with ChatGPT's performance in medical examinations (10.2%), specifically its ability to provide accurate answers to medical exam questions.

In contrast with the above, the main areas of research that are emerging or insufficiently studied are (1) the discipline-specific application of LLMs (e.g., in economics, social sciences, etc.), and (2) multi-modal generative AI in education, which integrates text, sound and visual information in teaching and learning.

Additionally, the authors have identified, based on their experience in the educational system and knowledge of gen-AI technologies, a series of gaps and future research opportunities or challenges to explore. These are grouped into several dimensions, as shown in Table 8.

Table 8. Future research opportunities.

Gaps and Challenges	Open Questions
Pedagogical gaps	Effectiveness in learning outcomes: Generative AI can personalize learning, but is there limited long-term effectiveness compared with traditional teaching and evaluation methods?
Ethical and social gaps	Critical thinking: Does reliance on AI-generated content reduce students' ability to critically evaluate information? Can AI foster literacy among learners?
Computational challenges	Digital divide and accessibility: How can access to AI-driven educational tools be democratized in low-income/rural areas?
Psychological and cognitive challenges	Plagiarism and academic integrity: How can educational institutions integrate AI while maintaining academic integrity?
	Bias in AI-generated content: How can stereotypes in educational materials be avoided?
	Misleading explainability: How can we make AI more precise and transparent to understand how it generates responses?
	Human–AI collaboration in teaching: How can AI be used to enhance teachers' roles rather than replace them? How should educational institutions prepare students for AI-driven job markets?
	Cognitive load and over-reliance: Does excessive AI use lead to passive learning?
	Personalization without isolation: Can AI-driven education personalize learning for ethnic and religious groups without leading to social isolation? Does the degree of personalization contribute to greater isolation or deeper integration in our lives?

Therefore, the integration of generative AI into educational systems presents several challenges. The gaps identified in Table 8—ranging from pedagogical concerns about long-term learning effectiveness and critical thinking to ethical issues, such as the digital divide and academic integrity, as well as computational and psychological challenges—highlight a complex landscape that warrants further exploration. Future research could address these open questions by evaluating long-term outcomes, enhancing critical thinking, ensuring ethical integration, improving transparency and collaboration and balancing personalization and social integration.

5. Conclusions

The introduction discusses the rapid evolution of gen-AI and LLMs, highlighting key milestones such as the introduction of transformer models in 2017, GPT-3 in 2020 and GPT-4 in 2024. These advancements have transformed applications of gen-AI, especially in education, whereby LLMs enable adaptive learning and personalized guidance but also raise concerns about academic integrity, plagiarism and ethical challenges.

This research aims to provide a bibliometric analysis of publications at the intersection of gen-AI and education to identify trends, distributions, opinions, gaps and emerging topics. The analysis also evaluates the effectiveness of gen-AI in teaching and learning while addressing challenges like data collection, quality selection and the limitations

of quantitative metrics. Additionally, it focuses on ethical issues such as intellectual property, dependency on AI, biases in AI outputs and the impact on critical thinking and interpersonal skills. The controversies in education and gen-AI are discussed, including the risks of an over-reliance on AI tools, challenges to traditional teaching methods and societal concerns about the role of teachers and students' readiness for job markets requiring human-centric skills. Ethical concerns include accountability for AI-generated content, biases and the cultural shift in education due to AI integration.

The research defines four research questions: (1) determining major contributing institutions, (2) identifying main topics by LDA, (3) exploring insufficiently studied topics and (4) understanding sentiment and novelty in the application of gen-AI in education. These research questions aim to provide insights into the evolving landscape of gen-AI in education and to guide future research efforts.

The sentiment analysis (SA) of publication abstracts on gen-AI in education reveals several conclusions. For all the methods, the majority of the publication abstracts maintained, as expected, a neutral writing style, reflecting the objective tone typical of academic literature. However, positive sentiments were consistently more prevalent than negative sentiments, with positive words appearing in about 80% of publications and negative words in only about 20%. All methods demonstrated an upward trend in positive sentiment from 2023 to 2024. The custom algorithm showed a notable increase in positive sentiment scores (from 0.705 in 2023 to 0.816 in 2024), while VADER and TextBlob metrics also confirmed a similar rise in positivity. Negative sentiment has slightly decreased over time. The findings from the custom word-count-based method were corroborated by the VADER and TextBlob analyses: VADER indicated a predominance of neutral sentiment (average neutral score: ~0.86) with a positive sentiment score (0.11) significantly higher than a negative sentiment (0.025). The compound score, measuring overall sentiment, also showed an increase from 0.77 in 2023 to 0.81 in 2024. TextBlob showed a positive polarity (~0.09) and consistent subjectivity (~0.47), suggesting that the abstracts were slightly positive and moderately objective. Polarity increased slightly in 2024, while subjectivity remained stable. The trend of rising positivity over time was consistent across all three methods, confirming the robustness of the findings.

In terms of the topic modeling results using LDA, five topics were identified as follows: (1) gen-AI's impact in education and research, (2) ChatGPT as a tool for university students and teachers, (3) large language models and prompting in computing education, (4) applications of ChatGPT in patient education and (5) ChatGPT's performance in medical examinations.

Insufficiently addressed topics, as suggested by the LDA weights, are Topic (3)—LLMs and prompting in computing education—which needs more in-depth exploration of pedagogical strategies and tools and Topic (4)—applications of ChatGPT in patient education—which requires more focus on implementation, effectiveness and accessibility for diverse patient demographics. Despite its low LDA weight, Topic (5)—ChatGPT's performance in medical examinations—may not be part of this category, as it covers a highly specialized area of interest.

These topics are relatively broad, and more fine-grained analysis would be recommended in future research in order to identify emerging topics in more detail. However, these findings indicate that studies of domain-specific LLM-based applications can be extended to other educational fields besides computing, health and patient education.

Given the current strong research focus on text-based/NLP-oriented technologies in education, the field of multi-modal generative AI in education, combining text, audio and video to create virtual laboratories or simulations, can certainly be considered an under-represented research direction.

Personalization is another promising avenue, with AI systems capable of creating individualized learning experiences tailored to a student's progress, preferences and needs. Researchers could also investigate the effectiveness of adaptive assessments designed by AI compared with traditional methods.

AI's role as a peer or tutor in collaborative learning environments could also be explored, particularly in fostering peer-to-peer engagement or by simulating study groups for students with limited access to collaborators.

Cross-cultural and multilingual applications are essential for designing AI tools that generate culturally relevant content and that facilitate the teaching of lesser-known languages. Similarly, the gamification of learning using generative AI offers possibilities for creating interactive storytelling scenarios and dynamically adjusting the difficulty in educational games to enhance engagement.

Hybrid human–AI classrooms represent a transformative concept whereby the dynamics of co-teaching by educators and AI systems, including the impact on teacher–student relationships and classroom authority, can be studied. Long-term studies on learning outcomes facilitated by AI compared with the use of traditional methods could provide valuable insights into knowledge retention and effectiveness. Finally, generative AI could support institutional decision-making in education, offering predictive insights for enrollment trends and curriculum optimization.

Author Contributions: Conceptualization, I.Î., S.-V.O., A.B. and D.V.; Formal analysis, I.Î.; Methodology, S.-V.O. and A.B.; Software, D.V.; Validation, S.-V.O. and A.B.; Visualization, I.Î.; Writing—original draft, I.Î. and S.-V.O.; Writing—review and editing, A.B. and D.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Bucharest University of Economic Studies through the 2024 Continuous Scientific Research Training Project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is publicly available at: <https://github.com/iintorsureanu/education-and-gen-ai-publications>.

Acknowledgments: This research was supported by the Bucharest University of Economic Studies through the 2024 Continuous Scientific Research Training Project.

Conflicts of Interest: The authors report that there are no competing interests to declare.

References

1. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3058.
2. Sohail, S.S.; Farhat, F.; Himeur, Y.; Nadeem, M.; Madsen, D.Ø.; Singh, Y.; Atalla, S.; Mansoor, W. Decoding ChatGPT: A Taxonomy of Existing Research, Current Challenges, and Possible Future Directions. *J. King Saud Univ. Comput. Inf. Sci.* **2023**, *35*, 101675. [[CrossRef](#)]
3. OpenAI. Hello GPT-4o. Available online: <https://openai.com/index/hello-gpt-4o/> (accessed on 9 December 2024).
4. Bozkurt, A. Generative Artificial Intelligence (AI) Powered Conversational Educational Agents: The Inevitable Paradigm Shift. *Asian J. Distance Educ.* **2023**, *18*, 718.
5. Hang, C.N.; Tan, C.W.; Yu, P.D. MCQGen: A Large Language Model-Driven MCQ Generator for Personalized Learning. *IEEE Access* **2024**, *12*, 102261–102273. [[CrossRef](#)]
6. Wei, Y.; Zhang, R.; Zhang, J.; Qi, D.; Cui, W. Research on Intelligent Grading of Physics Problems Based on Large Language Models. *Educ. Sci.* **2025**, *15*, 116. [[CrossRef](#)]
7. Dwivedi, Y.K.; Kshetri, N.; Hughes, L.; Slade, E.L.; Jeyaraj, A.; Kar, A.K.; Baabdullah, A.M.; Koohang, A.; Raghavan, V.; Ahuja, M.; et al. "So What If ChatGPT Wrote It?" Multidisciplinary Perspectives on Opportunities, Challenges and Implications of Generative Conversational AI for Research, Practice and Policy. *Int. J. Inf. Manag.* **2023**, *71*, 102642. [[CrossRef](#)]

8. Lin, S.M.; Chung, H.H.; Chung, F.L.; Lan, Y.J. Concerns About Using ChatGPT in Education. In *Innovative Technologies and Learning*; LNCS; Springer: Cham, Switzerlad, 2023; Volume 14099, pp. 37–49.
9. Elkhatat, A.M.; Elsaied, K.; Almeer, S. Evaluating the Efficacy of AI Content Detection Tools in Differentiating between Human and AI-Generated Text. *Int. J. Educ. Integr.* **2023**, *19*, 1–16. [[CrossRef](#)]
10. Donthu, N.; Kumar, S.; Mukherjee, D.; Pandey, N.; Lim, W.M. How to Conduct a Bibliometric Analysis: An Overview and Guidelines. *J. Bus. Res.* **2021**, *133*, 285–296. [[CrossRef](#)]
11. Li, L.; Ma, Z.; Fan, L.; Lee, S.; Yu, H.; Hemphill, L. ChatGPT in Education: A Discourse Analysis of Worries and Concerns on Social Media. *Educ. Inf. Technol.* **2024**, *29*, 10729–10762. [[CrossRef](#)]
12. Yan, L.; Sha, L.; Zhao, L.; Li, Y.; Martinez-Maldonado, R.; Chen, G.; Li, X.; Jin, Y.; Gašević, D. Practical and Ethical Challenges of Large Language Models in Education: A Systematic Scoping Review. *Br. J. Educ. Technol.* **2024**, *55*, 90–112. [[CrossRef](#)]
13. Perkins, M. Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era: ChatGPT and Beyond. *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 1–24. [[CrossRef](#)]
14. Tomlinson, B.T.; Torrance, A.W.; Black, R. ChatGPT and Works Scholarly: Best Practices and Legal Pitfalls in Writing with AI. *SMU L. Rev. F.* **2023**, *76*, 108. [[CrossRef](#)]
15. Darvishi, A.; Khosravi, H.; Sadiq, S.; Gašević, D.; Siemens, G. Impact of AI Assistance on Student Agency. *Comput. Educ.* **2024**, *210*, 104967. [[CrossRef](#)]
16. Zhai, X. Transforming Teachers’ Roles and Agencies in the Era of Generative AI: Perceptions, Acceptance, Knowledge, and Practices. *J. Sci. Educ. Technol.* **2024**, *2024*, 1–11. [[CrossRef](#)]
17. Chan, C.K.Y.; Tsi, L.H.Y. The AI Revolution in Education: Will AI Replace or Assist Teachers in Higher Education? *arXiv* **2023**. [[CrossRef](#)]
18. Mishara, P. The Ethical Implications of AI in Education: Privacy, Bias, and Accountability. *J. Inform. Educ. Res.* **2024**, *4*, 3550. [[CrossRef](#)]
19. Yusuf, A.; Pervin, N.; Román-González, M. Generative AI and the Future of Higher Education: A Threat to Academic Integrity or Reformation? Evidence from Multicultural Perspectives. *Int. J. Educ. Technol. High. Educ.* **2024**, *21*, 1–29. [[CrossRef](#)]
20. Ivanova, M.; Grosseck, G.; Holotescu, C. Unveiling Insights: A Bibliometric Analysis of Artificial Intelligence in Teaching. *Informatics* **2024**, *11*, 10. [[CrossRef](#)]
21. Qin, Q.; Zhang, S. Visualizing the Knowledge Mapping of Artificial Intelligence in Education: A Systematic Review. *Educ. Inf. Technol.* **2024**, *30*, 449–483. [[CrossRef](#)]
22. Pradana, M.; Elisa, H.P.; Syarifuddin, S. Discussing ChatGPT in Education: A Literature Review and Bibliometric Analysis. *Cogent Educ.* **2023**, *10*, 2243134. [[CrossRef](#)]
23. Bahroun, Z.; Anane, C.; Ahmed, V.; Zacca, A. Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. *Sustainability* **2023**, *15*, 12983. [[CrossRef](#)]
24. Yun, W.S.; Surianshah, S. A Review of the Scholarly Works on ChatGPT Use in Education: Bibliometric Analysis. *Int. J. Technol. Educ.* **2024**, *7*, 650–666. [[CrossRef](#)]
25. Wong, L.H.; Park, H.; Looi, C.K. From Hype to Insight: Exploring ChatGPT’s Early Footprint in Education via Altmetrics and Bibliometrics. *J. Comput. Assist. Learn.* **2024**, *40*, 1428–1446. [[CrossRef](#)]
26. Bozkurt, A. Unleashing the Potential of Generative AI, Conversational Agents and Chatbots in Educational Praxis: A Systematic Review and Bibliometric Analysis of GenAI in Education. *Open Prax.* **2023**, *15*, 261–270. [[CrossRef](#)]
27. Bhullar, P.S.; Joshi, M.; Chugh, R. ChatGPT in Higher Education—A Synthesis of the Literature and a Future Research Agenda. *Educ. Inf. Technol.* **2024**, *29*, 21501–21522. [[CrossRef](#)]
28. Saúde, S.; Barros, J.P.; Almeida, I. Impacts of Generative Artificial Intelligence in Higher Education: Research Trends and Students’ Perceptions. *Soc. Sci.* **2024**, *13*, 410. [[CrossRef](#)]
29. Watrianthos, R.; Ahmad, S.T.; Muskhir, M. Charting the Growth and Structure of Early ChatGPT-Education Research: A Bibliometric Study. *J. Inf. Technol. Educ. Innov. Pract.* **2023**, *22*, 235–253. [[CrossRef](#)]
30. Mahrishi, M.; Abbas, A.; Radovanovic, D.; Hosseini, S. Emerging Dynamics of ChatGPT in Academia: A Scoping Review. *J. Univ. Teach. Learn. Pract.* **2024**, *21*, 146–176. [[CrossRef](#)]
31. Chang, C.Y.; Chen, I.H.; Tang, K.Y. Roles and Research Trends of ChatGPT-Based Learning: A Bibliometric Analysis and Systematic Review. *Educ. Technol. Soc.* **2024**, *27*, 471–486. [[CrossRef](#)]
32. Zheltukhina, M.R.; Sergeeva, O.V.; Masalimova, A.R.; Budkevich, R.L.; Kosarenko, N.N.; Nesterov, G.V. A Bibliometric Analysis of Publications on ChatGPT in Education: Research Patterns and Topics. *Online J. Commun. Media Technol.* **2024**, *14*, e202405. [[CrossRef](#)]
33. Polat, H.; Topuz, A.C.; Yıldız, M.; Taşlibeyaz, E.; Kurşun, E. A Bibliometric Analysis of Research on ChatGPT in Education. *Int. J. Technol. Educ.* **2024**, *7*, 59–85. [[CrossRef](#)]
34. Amarathunga, B. ChatGPT in Education: Unveiling Frontiers and Future Directions through Systematic Literature Review and Bibliometric Analysis. *Asian Educ. Dev. Stud.* **2024**, *13*, 412–431. [[CrossRef](#)]

35. Grassini, S. Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Educ. Sci.* **2023**, *13*, 692. [[CrossRef](#)]
36. İpek, Z.H.; Gözüm, A.İ.C.; Papadakis, S.; Kallogiannakis, M. Educational Applications of the ChatGPT AI System: A Systematic Review Research. *Educ. Process Int. J.* **2023**, *12*, 26–55. [[CrossRef](#)]
37. Zhang, P.; Tur, G. A Systematic Review of ChatGPT Use in K-12 Education. *Eur. J. Educ.* **2024**, *59*, e12599. [[CrossRef](#)]
38. Ansari, A.N.; Ahmad, S.; Bhutta, S.M. Mapping the Global Evidence around the Use of ChatGPT in Higher Education: A Systematic Scoping Review. *Educ. Inf. Technol.* **2024**, *29*, 11281–11321. [[CrossRef](#)]
39. Peláez-Sánchez, I.C.; Velarde-Camaqui, D.; Glasserman-Morales, L.D. The Impact of Large Language Models on Higher Education: Exploring the Connection between AI and Education 4.0. *Front Educ.* **2024**, *9*, 1392091. [[CrossRef](#)]
40. Dempere, J.; Modugu, K.; Hesham, A.; Ramasamy, L.K. The Impact of ChatGPT on Higher Education. *Front Educ.* **2023**, *8*, 1206936. [[CrossRef](#)]
41. Hutto, C.J.; Gilbert, E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In Proceedings of the International AAAI Conference on Web and Social Media, Cologne, Germany, 16 May 2014; Volume 8, pp. 216–225.
42. Loria, S. TextBlob: Simplified Text Processing. Available online: <https://textblob.readthedocs.io/en/dev/index.html> (accessed on 5 December 2024).
43. Rehurek, R.; Sojka, P. Software Framework for Topic Modelling with Large Corpora. In Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, Valletta, Malta, 22 May 2010; pp. 45–50.
44. Röder, M.; Both, A.; Hinneburg, A. Exploring the Space of Topic Coherence Measures. In Proceedings of the 8th ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 399–408.
45. Sievert, C.; Shirley, K.E. LDAvis: A Method for Visualizing and Interpreting Topics. In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Baltimore, MD, USA, 17 June 2014; pp. 63–70.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.