Multidisciplinary Reviews

BIBLIOMETRIC RESEARCH
Published Online: April 10, 2026
https://doi.org/10.31893/multirev.2026524

ChatGPT in research and practice: Insights into applications, challenges, and future prospects

Ton Nguyen Trong Hien[a] | **Lislee Valle**[bf] ✉ | **Khoiriya Latifah**[cf] | **Muchamad Taufiq Anwar**[df] | **Rahul Bhandari**[e]

[a]Faculty of Business Administration, Van Lang University, Ho Chi Minh City, Vietnam.
[b]College of Education, Arts and Sciences, Cebu Technological University, Danao City, Cebu, Philippines.
[c]Faculty of Informatics and Engineering, Persatuan Guru Republik Indonesia, Semarang, Indonesia.
[d]Automotive Industry Information System, Politeknik STMI Jakarta, Central Jakarta, Indonesia.
[e]School of Business and International Office, O.P. Jindal Global University, Haryana, India.
[f]Naveen Jindal Young Global Research Fellowship, O.P. Jindal Global University, Haryana, India.

Abstract The emergence of ChatGPT, a sophisticated language model developed by OpenAI, has revolutionized the fields of artificial intelligence (AI) and natural language processing (NLP). This study offers a comprehensive overview of current research directions related to ChatGPT, analyzing its applications, challenges, and prospects. This most up-to-date study extracts and analyzes keywords and abstracts from 7,455 research articles in the Scopus database, providing a thorough overview using text mining techniques. The evaluation, conducted through descriptive statistical analysis of keyword frequency, co-occurrence, and related metrics, highlights current research issues, indicating that efficacy evaluation remains an ongoing area of investigation. Distinct from previous studies, this research extends the findings by classifying the overarching impact of ChatGPT across all fields, leveraging deep learning algorithms through Generative Artificial Intelligence (Gen AI), utilizing BERT Base Multilingual Uncased for sentiment analysis. Concurrently, it contrasts these results with those obtained from traditional analysis using the Vader Lexicon machine learning approach, revealing a consistent and overwhelmingly positive impact that significantly outweighs negative biases. While ChatGPT demonstrates transformative potential across diverse sectors, ranging from enhancing business operational strategies and content innovation to streamlining programming tasks and fostering personalized learning or clinical simulations, results also underscore critical concerns regarding academic integrity, biased AI responses, and diminished critical thinking. The metrics evaluation using Gen AI and machine learning shows good performance; The findings further suggest that Gen AI opens more promising avenues for enhancing bibliometric analysis, providing a robust framework for evaluating academic documents specifically, and advancing natural language processing methodologies more broadly.

Keywords: impact research, gen AI, machine learning, text mining

1. Introduction

ChatGPT, an abbreviation for Chat Generative Pre-Training Transformer, is a cutting-edge AI system engineered by OpenAI. Over the past two years, ChatGPT has become a widely recognized name for its ability to help users quickly and accurately search for information and answer inquiries. Designed to simulate human-like interaction, ChatGPT can engage in conversation and provide in-depth, easily understandable responses across various fields. It is therefore unsurprising that search queries related to ChatGPT and AI in general have surged since 2022.

The rapid adoption of ChatGPT can be attributed to its foundation in deep learning and natural language processing (NLP) technologies (Alomari, 2024; Javaid et al., 2023). By leveraging a vast dataset, it generates coherent, contextually relevant responses, making it suitable for a range of applications, including education, customer support, content creation, and research assistance (Mbwambo & Kaaya, 2024; Wang et al., 2024). Its ability to learn from patterns and predict user intent enables highly personalized interactions, which enhances user satisfaction. Furthermore, ChatGPT's flexibility in handling multiple languages and complex inquiries has expanded its global reach, addressing both personal and professional needs.

The rising interest in ChatGPT also reflects a broader trend in AI development, in which machine learning models are increasingly integrated into daily life. Industries are beginning to harness their potential to optimize workflows, automate tasks, and innovate service delivery. As AI tools like ChatGPT continue to evolve, ethical considerations such as data privacy, bias, and transparency remain critical challenges. Addressing these concerns will be essential in fostering trust and ensuring the responsible deployment of AI in the future. With its ongoing advancements, ChatGPT is poised to remain at the forefront of transformative technologies, driving productivity and reshaping human-computer interactions. The rise of ChatGPT has sparked a notable surge in interest in AI.

 Received: January 4, 2026 | Accepted: February 9, 2026

Technological advancements throughout history have consistently reshaped social relationships in profound ways. ChatGPT, in particular, merits attention for its wide range of applications, from offering technical assistance to addressing complex challenges. With its natural and efficient communication style, ChatGPT creates an experience similar to engaging in a real conversation, helping users feel understood and well-supported (Hu & Škultéty, 2024). However, like any major technological breakthrough, it has sparked discussions about the advantages and potential drawbacks of scientific progress. ChatGPT's rapid development has generated a variety of opinions within the community. Some see it as bringing numerous societal benefits (Baldassarre et al., 2023; Haque, 2024), while other studies argue it has minimal impact on business performance and primarily satisfies employees (Chu, 2023). Concerns arise about potential risks, such as misuse and over-reliance on AI, which could negatively affect various fields (Baldassarre et al., 2023; Haque, 2024). This study aims to provide a comprehensive view of ChatGPT's current issues and evaluate whether this attention is positive, negative, or neutral. The study uses text mining techniques to extract and analyze textual data, uncovering patterns, trends, and relationships within the information. By processing large volumes of text, the research identifies common themes, keywords, and concepts in the literature, uncovering patterns that might be challenging to discern manually (Jim et al., 2024). Although previous studies have explored the impact of ChatGPT using a bibliometric approach (Safdar et al., 2024; Khan et al., 2024), a consensus regarding its holistic impact remains elusive due to the fragmented nature of current academic contributions. Conventional bibliometric methods predominantly rely on keyword co-occurrence analysis; however, these approaches are inherently limited as they cannot interpret the underlying scholarly sentiment or the complex evaluative tones present in research findings (Alsolbi et al., 2022). In response to these gaps, this study adopts a dual-model framework designed to enhance the reliability of the findings through a rigorous comparative analysis. This study further employs machine learning models, using a rigorous comparative approach to capture context and the relationships between words within sentences, rather than relying solely on keyword frequency or co-occurrence, as seen in prior co-occurrence analyses. Consequently, the research aims to achieve three measurable objectives:

a. To map the global distribution and thematic evolution of ChatGPT research across the Scopus database.

b. To quantify the prevailing sentiment of academic abstracts, thereby identifying core scholarly attitudes toward AI integration.

c. To validate the reliability of Gen AI-based sentiment analysis through a comparative performance assessment against traditional machine learning lexicons, establishing its efficacy as a robust methodological tool.

2. Design/Methodology

The proposed method employs a bibliometric approach to search for research materials. This study employs a multi-stage analytical approach that combines descriptive statistics and sentiment classification analysis (Figure 1).

2.1. Literature search

The study utilized the Scopus database, a widely recognized and comprehensive source of peer-reviewed literature. Scopus was chosen for its extensive coverage, encompassing over 60% of the Web of Science database (Zhao & Strotmann, 2015), and its inclusion of diverse academic disciplines, ensuring a broad representation of relevant studies. Compared to other databases, Scopus offers an effective balance of breadth and quality, making it highly suitable for systematic reviews and bibliometric analysis.

To maintain the rigor and relevance of the study, the search process was carefully refined to include only articles published in journals and conference proceedings, as these sources are typically subjected to rigorous peer-review processes, ensuring their reliability and credibility. The search was further restricted to publications in English, as it remains the dominant language in scientific discourse. This approach allows access to a global body of research and ensures consistency in the analysis. Specifically, the search criteria included titles containing 'ChatGPT', limited to final-stage publications categorized as research articles or conference papers and published in journals or conference proceedings.

2.2. Article extraction

In this phase, all keywords and abstracts were extracted into a CSV file to prepare for the subsequent analysis step. Keywords and abstracts offer a concise view of the research's essence, making them valuable for understanding and interpreting the study's focus and impact. Specifically, the data were retrieved from the Scopus database, focusing exclusively on final articles and conference papers published in English. To ensure the integrity of the subsequent analysis, a rigorous preprocessing procedure was applied to the extracted text. This involved converting all content to lowercase, removing punctuation, and omitting non-informative stop words such as 'am', 'is', and 'are'. These refinement steps were essential to eliminate noise and standardize the dataset, thereby providing a clean foundation for the final analytical phase.

2.3. Data analysis

Initially, descriptive statistical methods were employed to present the characteristics and origins of the dataset. Subsequently, keyword analysis used frequency-based techniques to identify dominant themes within the textual data. For the sentiment analysis phase, two distinct approaches were implemented.

As the initial step in the classification analysis, a Gen AI model, specifically BERT Base Multilingual Uncased, was used to perform context-aware sentiment classification. This methodology has been applied in recent years to natural language analysis, including examining scientific research articles (Sharma et al., 2025). Introduced in 2019, this model has demonstrated remarkable effectiveness in numerous NLP tasks, leading to broad acknowledgment (Aziz et al., 2024).

In parallel, a second approach based on traditional lexicon-driven machine learning techniques, namely, Naïve Bayes, Random Forest, and Recurrent Neural Network (RNNs), was applied for sentiment categorization. Applying machine learning techniques, including these algorithms with TF-IDF vectorization, represents a well-established and reliable approach in traditional computational methods (Parida et al., 2021). Results derived from both approaches were compared and cross-validated against each other, allowing for a more comprehensive and trustworthy interpretation of sentiment across the dataset.

All analytical procedures were programmed and executed using the Python programming language.

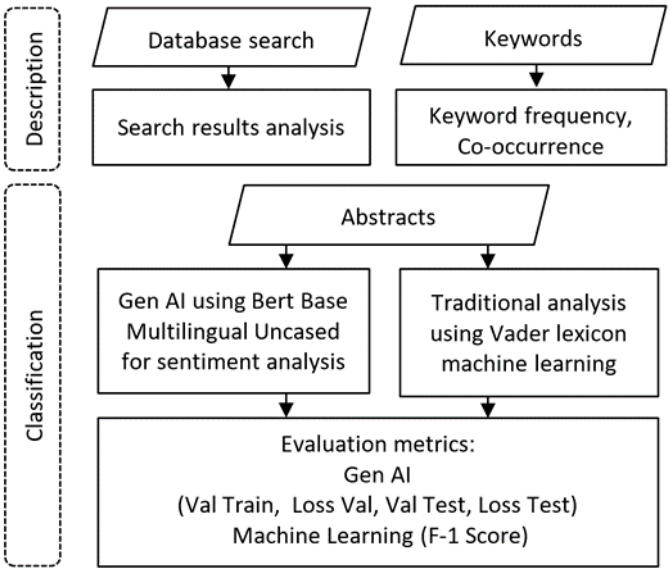

Figure 1 Detailed flow diagram of analysis.

3. Results and Discussion

3.1. Descriptive analysis

The study yielded a total of 10,909 documents through a search. After filtering, 7,455 documents met the specified criteria at the time of extraction. Among these, only 4 articles were published in 2022. In contrast, there were 1,618 articles in 2023; 3,192 in 2024, and 2,600 in 2025. The trends in this publication demonstrate a steep upward trajectory, with the unprecedented academic use of ChatGPT across multiple disciplines. The descriptive statistics (Figure 2) obtained are as follows: Computer Science emerged as the field with the highest proportion, accounting for 23.3%. This is anticipated, given that ChatGPT is an application within artificial intelligence and NLP, both of which are integral to Computer Science. Following this, Social Sciences ranked second with 19.3%, reflecting considerable interest in studying ChatGPT from a societal perspective, such as its impact on society, media, and education. Medicine occupied the third position with 15.5%, indicating the use of ChatGPT in medical support, information provision, and patient communication. Additionally, fields such as engineering (8.6%) and mathematics (5.5%) also make significant contributions. Collectively, these fields demonstrate the diverse research landscape surrounding ChatGPT.

In terms of authorship distribution, analyzing author origins contributes to a deeper understanding of the global research landscape and can guide future research strategies and collaborations. The results of this study offer insights into the diversity of the research community. The majority of authors are affiliated with institutions in the United States (1,824 articles), followed by China (961 articles) and Germany (417 articles). Turkey (401) and India (99) also make notable contributions. It is noteworthy that the leading contributors are technologically advanced nations, such as the United States and China; however, emerging economies like India and Turkey are beginning to adopt ChatGPT research.

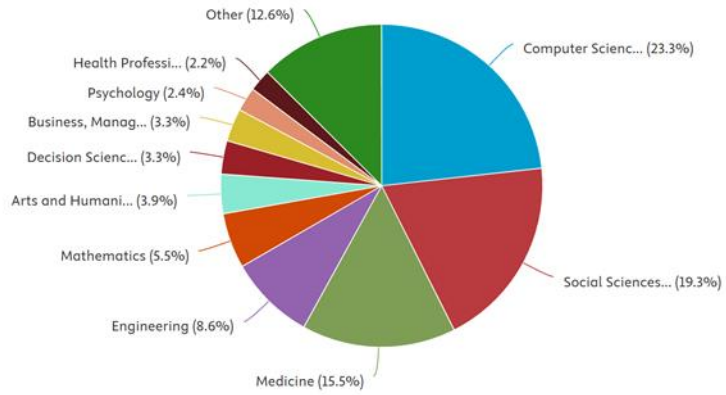

Figure 2 The diagram illustrates the distribution of research fields. *Source:* Scopus Database.

ChatGPT, AI, and Large Language Model (LLM) are core, foundational concepts within the field of computer science, particularly its subfields of artificial intelligence and machine learning. Consequently, the high frequency of these keywords is unsurprising. Avoid aggregating common nouns such as ChatGPT, artificial intelligence, 'chatbot', 'large language model', etc. These common nouns appear frequently and may dominate the analysis, making it challenging to identify specific research topics or issues. Excluding these keywords, we can focus on other keywords and subtopics that may indicate more specific research areas, thereby preventing these general terms from diluting the data and reducing the ability to detect significant trends or issues. The results of the analysis are summarized in Figure 3.

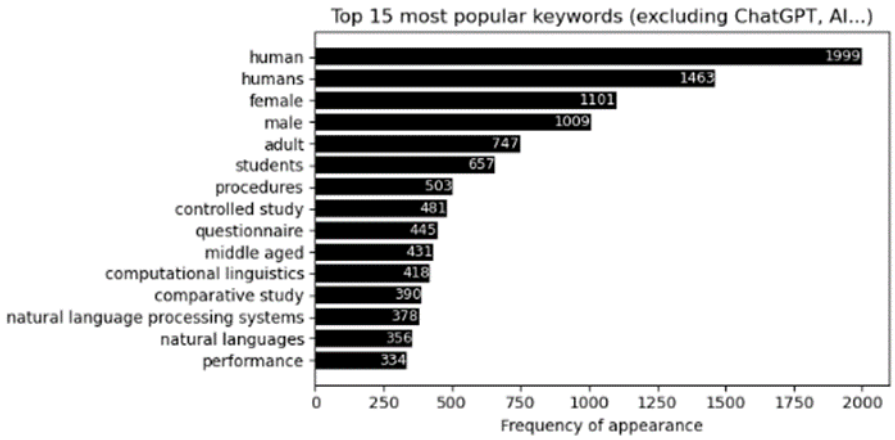

Figure 3 Top 15 most popular keywords (excluding ChatGPT, AI, etc..). *Source:* Scopus Database.

A keyword frequency analysis Figure 3 reveals clear trends in application and academic interest, with a strong pivot toward human-centric evaluation and demographic-centric assessment. Notably, the keywords 'human' (1,999 occurrences). 'humans' (1,463 occurrences), 'female' (1,101), 'male' (1,009), 'adult' (747), or 'middle aged' (431) overwhelmingly dominate the list, indicating that a substantial portion of current research focuses on human interaction and relationships with large language models. Scholars are examining demographic variations, biases, and the applicability of ChatGPT across various user groups, confirming concerns about the model's impact on diverse populations and settings.

Furthermore, the continued prominence of keywords pointing toward academic and controlled assessment remains vital. 'Students' (657) continues to emphasize the education sector, while 'procedures' (503), 'controlled study' (481), and 'questionnaire' (445) reflect a strong methodological effort toward rigorous empirical assessment. These terms indicate an ongoing focus on evaluating the quality, reliability, and accuracy of ChatGPT as a pedagogical tool. The new keyword 'performance' (334) further reinforces this critical evaluation focus, emphasizing the need to quantify the model's efficacy across tasks. This implies that existing research on ChatGPT in education primarily focuses on student learning behaviors, academic integrity, performance assessment, and the integration of ChatGPT into instructional and assessment tasks.

The foundational technical and linguistic aspects are also maintained through keywords like 'computational linguistics' (418), 'natural language processing systems' (378), and 'natural languages' (356). This sustained focus on the underlying linguistic behavior and technical frameworks ensures a deeper understanding of the models, which is essential for effective integration across application domains. Finally, the inclusion of 'comparative study' (390) signals the continued trend of multi-model evaluations aimed at assessing tool efficacy and identifying best practices across specific application domains. This further implies that these trends have shifted from early exploratory studies toward empirical evaluation, domain-specific testing, and multi-model comparisons.

The analysis of keyword co-occurrence networks (Figure 4) reveals prominent thematic clusters that reinforce prior assertions when providing deeper insights into academic research directions. The central cluster, comprising keywords such as 'higher education', 'students', 'assessment', and 'educational technology', underscores a profound interest in the application of ChatGPT for educational innovation, particularly in higher education. Associations with keywords such as 'technology acceptance', 'UTAUT2' (Unified Theory of Acceptance and Use of Technology 2), and 'academic integrity' suggest a sustained research focus on how learners adopt and interact with ChatGPT. The keywords show that studies extend beyond technological integration in education, there is a focus on acceptance levels, perceived utility, trust, ethical risks, and the reshaping of assessment practices within educational contexts. This supports the need to balance pedagogical benefits with the risks associated with plagiarism and overreliance on using Gen AI tools like ChatGPT in education.

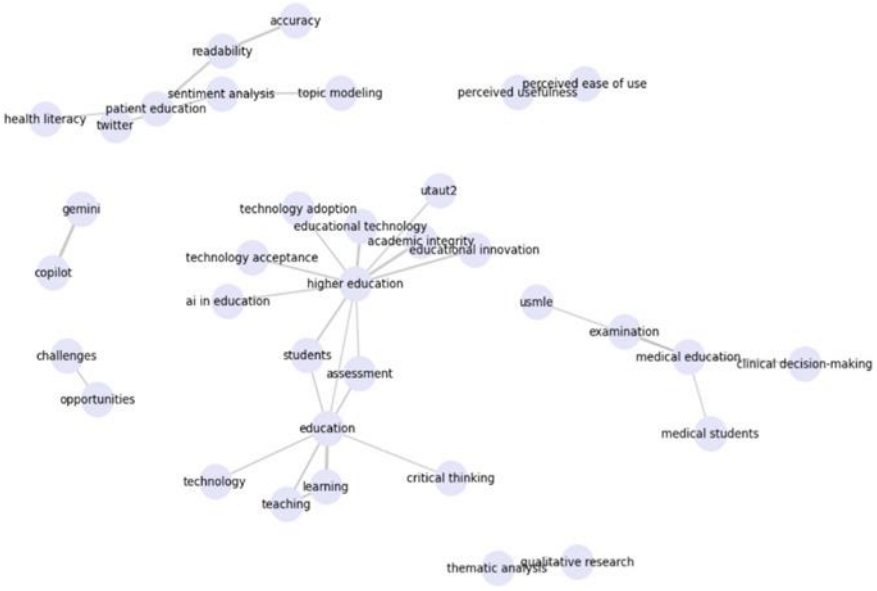

Figure 4 Co-occurrence keywords analysis. *Source:* Scopus Database.

Notably, the keyword cluster, including 'medical education', 'clinical decision-making', 'medical students', and 'USMLE' (United States Medical Licensing Examination), highlights ChatGPT's potential as a tool for training medical competencies, with the ability to simulate clinical processes and support intensive exam preparation. Concurrently, another group of keywords, such as 'sentiment analysis', 'readability', and 'topic modeling', reflects a research trajectory evaluating the quality of ChatGPT's outputs and its linguistic expressiveness, particularly in health communication applications like 'patient education'.

A significant additional finding is the linkage between the keywords 'opportunities' and 'challenges', suggesting that studies not only extol ChatGPT's potential but also critically address its barriers and risks. The presence of this keyword pair signals balanced research approaches, wherein the benefits and limitations of the technology are compared for a more comprehensive evaluation. Furthermore, the co-occurrence of 'qualitative research' and 'thematic analysis' indicates an emerging wave of qualitative studies aimed at exploring how users, students, or professionals interact with and perceive ChatGPT. This is particularly crucial, as earlier evaluations often leaned toward quantitative or technical assessments, whereas qualitative research offers clearer insights into user behavior, emotions, and the contextual realities of engaging with this large language model.

3.2. Classification analysis

This study evaluated two different models of sentiment analysis on the Scopus abstract dataset, namely Gen AI using Transfer Learning Bert for sequence classification (Bert-base-multilingual-uncased-sentiment) model, and machine learning traditional model classification.

Sentiment results using Bert-base-multilingual-uncased-sentiment were obtained, with a composition of 5,495 data points having positive sentiment, 1,458 having neutral sentiment, and 502 having negative sentiment. Positive sentiment is obtained if the label is 4 or 5 stars, neutral sentiment if the label is 3 stars, and negative sentiment if the label is 2 or 3 stars.

The measurement or evaluation matrix of the model has accuracy, precision, recall, and F-1 score values, all of which are 1, which means that the model has a good performance based on real-time testing, and appropriate sentiments are obtained, such as the following example: GPT is the best model, so the classification result is positive.

The second methodology employed in this study involves the application of traditional machine learning techniques for text analysis. This approach utilizes the VADER (Valence Aware Dictionary and Sentiment Reasoner) lexicon method, which assigns numeric labels to sentiment categories: -1 for negative, 0 for neutral, and 1 for positive, based on the VADER Sentiment

Analysis framework. The traditional machine learning process incorporates label generation from the dataset using the SentimentIntensityAnalyzer, a component of the VADER Sentiment library. VADER is a specialized tool designed for sentiment analysis, particularly effective for social media texts and unstructured data.

The subsequent step involves initializing an instance of the SentimentIntensityAnalyzer class using the VADER lexicon. This instance facilitates the sentiment analysis of textual data. A function, termed vader_labeling, is defined to accept a text parameter, enabling the classification of sentiment within the input text. This function invokes the polarity_scores() method to evaluate the sentiment, yielding a dictionary containing four scores: 'neg' (negative), 'neu' (neutral), 'pos' (positive), and 'compound'. The 'compound' score, which normalizes the overall sentiment to a range between -1 and +1, serves as the basis for classification. Sentiment is categorized as follows: a compound score ≥ 0.05 is classified as positive (returning 1), ≤ -0.05 as negative (returning -1), and values between -0.05 and 0.05 as neutral (returning 0).

The overall output of the VADER lexicon sentiment analysis process is 7455. The analysis shows a predominance of positive sentiment (6,557 instances) with fewer negative (562) and neutral (336) sentiments, indicating a generally optimistic view in the evaluated texts.

Following the labeling process utilizing the VADER lexicon, feature extraction was performed using the Term Frequency-Inverse Document Frequency (TF-IDF) method. Subsequently, classification techniques were applied to the outcomes derived from the feature extraction process. The results, presented in Table 1, obtained through the implementation of Naive Bayes, Random Forest, and RNNs are characterized by their precision, recall, and accuracy, as detailed in the table below:

Table 1 Performance metrics of sentiment analysis models.

	Labeling and Feature Extraction	Accuracy	Precision	Recall	F-1 Score
Naïve Bayes	Lexicon, TF-IDF	0.92	0.92	0.94	0.94
Random Forest	Lexicon, TF-IDF	0.92	0.92	0.89	0.92
Gen AI (Bert Base Multilingual Uncased)	BERT	1	1	1	1
	Labeling and Feature Extraction	Train Val	Loss Val	Train Test	Loss Test
RNNs	Lexicon, TF-IDF	0,89	0,41	0,89	0,40

An analysis of the presented data (Figure 5) reveals that the Gen AI approach (A), potentially linked to the BERT model, exhibits a pronounced tendency to classify sentiments as positive, with over 5,550 abstracts categorized as such. In comparison, the number of neutral abstracts (over 1,200) and negative abstracts (under 500) is considerably lower. This may reflect Gen AI's capability to capture complex semantic features and linguistic contexts. In contrast, the traditional Vader method (B) demonstrates a more balanced distribution, with approximately 5,100 abstracts classified as positive and fewer than 500 as negative, aligning with its lexicon-based approach, which typically relies on fixed semantic rules.

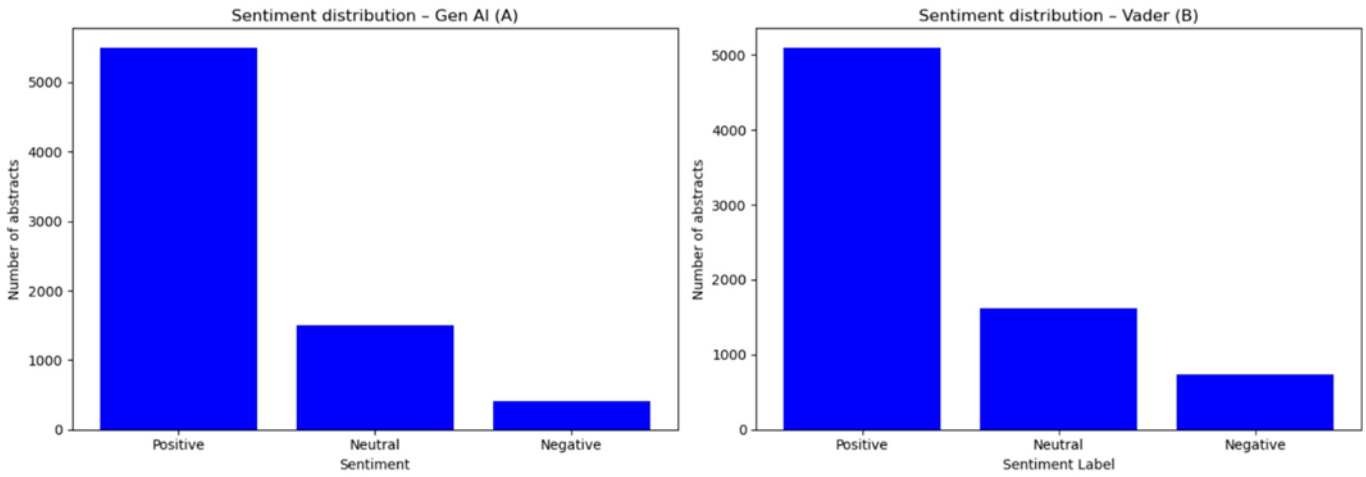

Figure 5 Gen AI sentiment analysis (A) vs. traditional analysis using the Vader lexicon (B). *Source:* Scopus Database.

Overall, the application of ChatGPT demonstrates a predominantly positive trend across the domains under investigation. The research aligns with the findings of Lo (2023), Ray (2023), and Wu et al. (2023), which suggest that ChatGPT's impact exerts a favorable influence across various fields, particularly highlighting how positive outcomes are reshaping the operational strategies and customer interactions of businesses, yielding significant economic benefits such as content innovation and enhanced market analysis efficiency. Furthermore, ChatGPT is capable of generating code snippets or even complete scripts based on natural language descriptions, thereby relieving programmers from the repetitive task of rewriting common code segments. In the educational sphere, Pradana et al. (2023) and Albadarin et al. (2024) acknowledge ChatGPT's role in facilitating personalized learning, while Montenegro-Rueda et al. (2023) emphasize its contribution to improving

teaching and learning processes. Students predominantly employ ChatGPT as a virtual intelligent assistant, providing immediate feedback, on-demand responses, and elucidations of complex subjects. Additionally, it serves as a tool to refine writing and linguistic proficiency through activities such as brainstorming, essay drafting, summarization, translation, rephrasing, and grammar correction, among others (Pradana et al., 2023; Albadarin et al., 2024). ChatGPT can significantly benefit learning when used responsibly and appropriately; however, in the long term, excessive reliance on ChatGPT may lead to issues such as diminished critical thinking, reduced academic motivation, and diminished social interaction (Lo et al., 2024).

Although negative indicators are present, their prevalence remains minimal compared to positive outcomes. Sadallah et al. (2025) highlight concerns regarding the use of ChatGPT in education. Despite its potential to enhance student engagement and access to materials, educators worry about its impact on learners' goals and stress levels. According to Nazir and Wang (2023) and Ali et al. (2024), key issues include biased AI responses, heightened plagiarism risks, and threats to authentic learning and independent thinking. These factors create pressure for students and educators, prompting institutions to strike a balance between technology use and academic integrity (Wu et al., 2023; Baig & Yadegaridehkordi, 2024).

Lo (2023) and Montenegro-Rueda et al. (2023) further argue that not only students but also educators require training to leverage ChatGPT effectively. Lo (2023) also notes the unsatisfactory performance of ChatGPT in mathematics. As a language model rather than a computational tool, ChatGPT does not execute mathematical calculations with logical precision but relies on patterns and correlations derived from its training data to predict the most plausible responses. This approach can yield inaccurate outcomes, particularly for complex or less frequently encountered computations. In the medical field, Ray (2023) and Li et al. (2024) assert that ChatGPT lacks reliability. The current version of ChatGPT has demonstrated only moderate or 'passing' performance across various medical assessments and applications; it is considered unsuitable for actual clinical implementation. This unreliability stems from its design, which is not optimized for critical clinical applications (Li et al., 2024). A point of convergence between this study and the prior work of Singh and Singh (2023) lies in the persistent uncertainty surrounding advanced technologies, particularly novel and continuously evolving products such as ChatGPT, which remains prevalent to date. Singh and Singh (2023) posited that inherent uncertainties within any technology or system may impact its performance and reliability.

4. Final Considerations

Research on ChatGPT has grown substantially since its inception, with notable contributions from researchers across the globe, especially in the United States and China. The results obtained demonstrate the widespread popularity and effectiveness of ChatGPT across various fields. While many studies focus on exploring the benefits of ChatGPT and its potential applications, the results also highlight the need for further research to address current challenges and ensure the effective and safe application of ChatGPT in the future. Specifically, the integration of AI support in education, including medical education, is highlighted as a primary concern among researchers, alongside worries about the erosion of critical thinking, accuracy in medical diagnostics, and other related issues. The analysis reveals that ongoing research continues to contribute to the development and validation of ChatGPT's accuracy. The study further suggests the efficacy of Gen AI and machine learning techniques in facilitating the assessment of academic publications. Future research should expand the evaluation to include additional databases and conduct empirical investigations to substantiate this theoretical contribution, thereby enhancing the robustness and generalizability of the findings. A notable limitation of this study is also the search scope, which was restricted to articles with 'ChatGPT' specifically in titles. While this ensured a high degree of thematic relevance, it may have excluded pertinent studies that discuss ChatGPT within abstracts or keywords under broader terms. Therefore, future research should expand the evaluation to include additional databases and broader search strings.

Acknowledgment

The paper acknowledges the participants and discussants from ICBESS 2025. This research also benefited from the kind support of O.P. Jindal Global University, which granted access to the Scopus database.

6. Declarations

6.1. Ethical considerations

Not applicable.

6.2. Use of artificial intelligence (AI)

The authors declare that the generative artificial intelligence (AI) tools, Grammarly and ChatGPT, were used exclusively for language editing and/or grammatical improvement. The use of AI did not influence the scientific content, study design, data analysis, data interpretation, results, or conclusions of the manuscript. Full responsibility for the content remains with the authors.

6.3. Conflict of interest

The authors declare no conflicts of interest. The authors collaboratively contributed diverse perspectives within the Naveen Jindal Young Global Research Fellowship.

6.4. Funding

This research did not receive any financial support.

References

Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education, 3*(1). https://doi.org/10.1007/s44217-024-00138-2

Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. *Education Sciences, 14*(6), 643. https://doi.org/10.3390/educsci14060643

Alomari, E. (2024). Unlocking the potential: A comprehensive systematic review of ChatGPT in natural language processing tasks. *Computer Modeling in Engineering & Sciences, 141*(1), 43.

Alsolbi, I., Wu, M., Zhang, Y., Joshi, S., Sharma, M., Tafavogh, S., Sinha, A., & Prasad, M. (2022). Different approaches of bibliometric analysis for data analytics applications in non-profit organisations. *Journal of Smart Environments and Green Computing, 2*(3), 90–104. https://doi.org/10.20517/jsegc.2022.09

Aziz, K., Ji, D., Chakrabarti, P., Chakrabarti, T., Iqbal, M. S., & Abbasi, R. (2024). Unifying aspect-based sentiment analysis BERT and multi-layered graph convolutional networks for comprehensive sentiment dissection. *Scientific Reports, 14*(1). https://doi.org/10.1038/s41598-024-61886-7

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in higher education: A systematic literature review and research challenges. *International Journal of Educational Research, 127*, 102411. https://doi.org/10.1016/j.ijer.2024.102411

Baldassarre, M. T., Caivano, D., Fernandez Nieto, B., Gigante, D., & Ragone, A. (2023, September). The social impact of generative AI: An analysis on ChatGPT. In *Proceedings of the 2023 ACM Conference on Information Technology for Social Good* (pp. 363–373).

Chu, M. N. (2023). Assessing the benefits of ChatGPT for business: An empirical study on organizational performance. *IEEE Access, 11*, 76427–76436. https://doi.org/10.1109/access.2023.3297447

Haque, M. A., & Li, S. (2024). Exploring ChatGPT and its impact on society. *AI and Ethics*. https://doi.org/10.1007/s43681-024-00435-4

Hu, W. C., & Škultéty, R. (2024). Unlocking the learning potential: ChatGPT as a virtual platform for cross-interaction in English language learning. *Engineering Proceedings, 74*(1), 59. https://doi.org/10.3390/engproc2024074059

Javaid, M., Haleem, A., & Singh, R. P. (2023). A study on ChatGPT for Industry 4.0: Background, potentials, challenges, and eventualities. *Journal of Economy and Technology, 1*, 127–143. https://doi.org/10.1016/j.ject.2023.08.001

Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, 100059. https://doi.org/10.1016/j.nlp.2024.100059

Jo, H., & Park, D. (2024). Effects of ChatGPT's AI capabilities and human-like traits on spreading information in work environments. *Scientific Reports, 14*, 7806. https://doi.org/10.1038/s41598-024-57977-0

Khan, N., Khan, Z., Koubaa, A., Khan, M. K., & Salleh, R. B. (2024). Global insights and the impact of generative AI–ChatGPT on multidisciplinary research: A systematic review and bibliometric analysis. *Connection Science, 36*(1). https://doi.org/10.1080/09540091.2024.2353630

Li, J., Dada, A., Puladi, B., Kleesiek, J., & Egger, J. (2024). ChatGPT in healthcare: A taxonomy and systematic review. *Computer Methods and Programs in Biomedicine, 245*, 108013. https://doi.org/10.1016/j.cmpb.2024.108013

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Lo, C. K., Hew, K. F., & Jong, M. S. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education, 219*, 105100. https://doi.org/10.1016/j.compedu.2024.105100

Mbwambo, N. M., & Kaaya, P. B. (2024). ChatGPT in education: Applications, concerns and recommendations. *Journal of ICT Systems, 2*(1), 107–124. https://doi.org/10.56279/jicts.v2i1.9020

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. *Computers, 12*(8), 153. https://doi.org/10.3390/computers12080153

Nazir, A., & Wang, Z. (2023). A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges. *Meta-Radiology, 1*(2), 100022. https://doi.org/10.1016/j.metrad.2023.100022

Parida, U., Nayak, M., & Nayak, A. K. (2021). News text categorization using random forest and naïve Bayes. In *Proceedings of the 2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)* (pp. 1–4). Bhubaneswar, India. https://doi.org/10.1109/ODICON50556.2021.9428925

Pradana, M., Elisa, H. P., & Syarifuddin, S. (2023). Discussing ChatGPT in education: A literature review and bibliometric analysis. *Cogent Education, 10*(2). https://doi.org/10.1080/2331186X.2023.2243134

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Sadallah, M., Bin-Nashwan, S. A., & Benlahcene, A. (2025). ChatGPT: A transformative role in academia—Insights into academic staff performance since adoption. *Journal of Information, Communication and Ethics in Society, 23*(1), 32–53. https://doi.org/10.1108/JICES-07-2024-0097

Safdar, M., Siddique, N., Gulzar, A., Adil, S. A., Yasin, H., & Khan, M. A. (2024). A bibliometric analysis of literature published on ChatGPT and GPT. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/gkmc-08-2023-0304

Sharma, R., Singh, B., & Khamparia, A. (2025). Machine learning and generative AI techniques for sentiment analysis with applications. In *Generative artificial intelligence for biomedical and smart health informatics* (pp. 183–208). https://doi.org/10.1002/9781394280735.ch10

Singh, H., & Singh, A. (2023). ChatGPT: Systematic review, applications, and agenda for multidisciplinary research. *Journal of Chinese Economic and Business Studies, 21*(2), 193–212. https://doi.org/10.1080/14765284.2023.2210482

Wang, B., Li, S., Dong, Y., & Zhang, H. (2024). ChatGPT-aided education teaching. In *Proceedings of the 2024 6th International Conference on Computer Science and Technologies in Education (CSTE)* (pp. 141–145). Xi'an, China. https://doi.org/10.1109/CSTE62025.2024.00033

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica, 10*(5), 1122–1136. https://doi.org/10.1109/jas.2023.123618

Zhao, D., & Strotmann, A. (2015). *Analysis and visualization of citation networks*. Morgan & Claypool Publishers. https://doi.org/10.1007/978-3-031-02291-3