

## Scientometric Analysis of ChatGPT

**Authors:** Yue Lei, Yue Lei

**Date:** 2023-04-11T00:00:00+00:00

### Abstract

**Purpose/Significance:** As a new generation of artificial intelligence, ChatGPT may exert influences on various aspects of social life. This study systematically reviews the current status of ChatGPT scientific research to provide references for future research endeavors. **Method/Process:** Through scientometric analysis methods, 161 ChatGPT-related documents included in CNKI were analyzed. **Results/Conclusions:** Current ChatGPT research is distributed across 40 different fields, with insufficient collaboration among scholars, and is primarily concentrated in coastal regions. The research primarily involves discussions on ChatGPT technology and its impact on various industries, particularly its influence on education. Current ChatGPT research remains in its early stages, necessitating more empirical studies to verify its actual impact in real-life contexts.

### Full Text

## Scientometrics Analysis of ChatGPT

Ziyang College of Environmental Science and Technology  
Ziyang, 641300

### Abstract:

**[Objective/Significance]** As a new generation of artificial intelligence, ChatGPT may impact various aspects of social life. This study reviews the current state of ChatGPT research to provide references for future related studies. **[Method/Process]** Using scientometric analysis methods, we analyzed 161 ChatGPT-related documents from CNKI. **[Results/Conclusion]** Current ChatGPT research is distributed across 40 different fields, with insufficient collaboration among scholars and a geographic concentration in coastal regions. The research primarily involves discussions of ChatGPT technology and its impact on various industries, particularly education. ChatGPT research remains

in its early stages, requiring more empirical studies to verify its practical impacts in real-world contexts.

**Keywords:** ChatGPT; scientometrics; R; VOSviewer

On November 30, 2022, OpenAI released ChatGPT, which reached one million registered users within just five days. Within two months, this figure soared to an astonishing 100 million users. This breakthrough sent ripples across the tech industry, prompting major internet companies both domestically and internationally—including Meta, Baidu, Google, and Alibaba—to rush to develop similar products. Concurrently, various industries began joining the discussion about ChatGPT, and its popularity in online media continued to rise (see Figure 1 [Figure 1: see original paper]).

ChatGPT (Chat Generative Pre-trained Transformer), known in Chinese as the “Generative Pre-trained Transformer Model,” is an artificial intelligence-driven natural language processing tool capable of completing tasks such as script writing, code generation, translation, and dialogue. Compared with traditional AI, ChatGPT features larger data sources, stronger computing power and algorithms, and a high degree of human-like intelligence [1]. On one hand, ChatGPT can improve work efficiency and optimize services [1-3]. On the other hand, its use also raises numerous risks and ethical considerations that must be addressed [4, 5].

As a blockbuster AI application, ChatGPT has also attracted widespread attention in China, with several journals already publishing papers on the topic. This study investigates the current state of ChatGPT research based on CNKI to provide information for understanding and studying ChatGPT.

**Figure 1.** Google Trends for ChatGPT from November 30, 2022 to April 2, 2023 [Figure 1: see original paper]. *Note: These numbers represent the popularity trend of search terms within a specific region and time range, using the highest point in the chart as the benchmark. A score of 100 indicates the term’s popularity peaked at that time, 50 indicates half the peak popularity, and 0 indicates insufficient data.*

## 2 Data and Methods

This study uses CNKI as the data source, conducting a subject search with “ChatGPT” as the search term on April 3, 2023. The search yielded 448 newspaper articles and 381 academic journal articles (161 of which were from the Chinese core journal database).

This study employs scientometric methods to conduct statistical analysis on the data from these 161 Chinese academic journal articles and uses software such as R and VOSviewer for visualization [6-8]. The analysis focuses on two aspects: (1) descriptive analysis of ChatGPT literature output, and (2) hotspot topic analysis of ChatGPT research.

### 3.1 Descriptive Statistics of ChatGPT Literature

In terms of research fields, according to CNKI data, ChatGPT research is distributed across 40 different domains. The fields with the highest paper output include automation technology (103 articles), information economy and postal economy (36 articles), computer software and applications (33 articles), educational theory and management (25 articles), journalism and media (12 articles), and library and information science (10 articles). As a new leading product in artificial intelligence capable of processing various language tasks, automation technology is currently the primary focus area.

The geographic distribution of ChatGPT research is shown in Figures 2A and 2B [Figure 2: see original paper]. Currently, 19 regions in China have published papers on ChatGPT. Overall, published research is concentrated in coastal provinces. Beijing ranks first with 77 articles, followed by Shanghai (24 articles) and Guangdong (10 articles).

Regarding author collaboration (as shown in Figure 2D), 65 articles have two or more authors, accounting for 40.34% of the total. Figure 2D reveals that the overall collaboration network is highly disconnected. As an emerging phenomenon, researchers' collaboration intensity in this field is somewhat insufficient, indicating strong potential for future collaboration.

Download frequency refers to the number of times a paper is saved or downloaded from journal websites, publishers, or databases—a product of digital publishing [9]. For research on emerging topics like ChatGPT, download counts may be more sensitive than citation counts. Analyzing the download frequency distribution of the 161 ChatGPT articles to identify highly downloaded papers reveals that over one-third have more than 1,000 downloads, indicating that ChatGPT has also attracted significant attention among Chinese researchers.

### 3.2 Hot Topic Analysis of ChatGPT Research

Keyword analysis of ChatGPT literature was conducted<sup>1</sup>, with results shown in Figure 3 [Figure 3: see original paper]. Three major clusters emerged: (1) ChatGPT technology, (2) ChatGPT's impact on education, and (3) ChatGPT's impact on other industries.

<sup>1</sup> Parameter settings for the mapping: minimum keyword co-occurrence frequency of 2, replacement of “gpt” with “chatgpt”, exclusion of disconnected keywords, and minimum cluster size of 4.

**Cluster 1** focuses on discussions of ChatGPT technology. Qian Li et al. mapped the overall technical architecture of ChatGPT and systematically analyzed its corpus system, pre-training algorithms and models, and fine-tuning algorithms and models [10]. Guo Quanzhong and Zhang Jinyi summarized four major technical features of ChatGPT (single modality, massive data, human annotation, and reinforcement learning) [11], while Zhu Guanghui and Wang Xiwen also analyzed its operational modes and key technologies [12].

**Cluster 2** addresses ChatGPT's impact on the education sector. Wu Xiaolin and Xing Yifei argue that as a new generation of AI, ChatGPT may challenge knowledge transmission-centered educational models, though its limitations make it unlikely to replace human teaching in the short term [13]. Zhong Binglin et al., through a written discussion, analyzed the challenges ChatGPT applications pose to higher education and their potential to drive educational transformation [14]. Shen Chao also examined how ChatGPT as a new technology can drive educational reform and cultivate innovative talent [15]. Jiao Jianli believes ChatGPT can facilitate the digital transformation of school education [16]. Wang Youmei et al. dialectically analyzed both the potential and risks of ChatGPT applications [17].

**Cluster 3** concerns ChatGPT's impact on industries beyond education. Zhang Xiaheng analyzed ChatGPT's influence on digital economic development, noting that it can promote digital transformation across multiple fields including education, healthcare, media, and e-commerce, enhance digital efficiency, broaden digital boundaries, and improve high-quality digital economic development [18]. Zhou Zhibo discussed the risks of introducing ChatGPT into digital government construction and how to regulate its application [19]. Researchers have also discussed its impact in fields such as social work, accounting, and information services [20-22].

## 4 Discussion

This study systematically analyzes the current state of domestic ChatGPT research from a scientometric perspective to provide an informational reference for future studies. The following conclusions are drawn: (1) Based on CNKI search results, ChatGPT, as a hot topic, currently shows limited overall publication activity, though new papers continue to emerge. Geographically, literature is mainly distributed in coastal regions, showing some consistency with regional economic development disparities. In terms of download counts, ChatGPT is widely recognized as a social hotspot. (2) Regarding research themes, ChatGPT studies primarily involve technical discussions and theoretical explorations of application impacts. In terms of field distribution, Chinese ChatGPT research spans 40 different domains, with researchers and practitioners across various industries beginning to notice ChatGPT's impact on their respective fields.

In summary, as an emerging phenomenon, ChatGPT research remains in its early stages. Current studies mainly focus on technical discussions and theoretical explorations of application impacts, while lacking substantial empirical research examining ChatGPT's real-world effects on various industries and individual aspects in practice. As a remarkably capable blockbuster AI, prohibition and avoidance are no longer viable options. Therefore, future researchers must rely on actual data to explore the impacts of ChatGPT-like AI on all aspects of social life, enabling the public to view the advantages and disadvantages of such AI dialectically.

## References

- [1] Zheng Shilin, Yao Shouyu, Wang Chunfeng. Economic and Social Impacts of ChatGPT New Generation AI Technology Development[J/OL]. *Industrial Economics Review* (2023-03-12). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyE>
- [2] van Dis E A M, Bollen J, Zuidema W, et al. ChatGPT: five priorities for research[J]. *Nature*, 2013, 614: 224-226.
- [3] Zhang Xiaoheng. New Generation AI Technology (ChatGPT) and Its Impact and Transformation on Human Society[J/OL]. *Industrial Economics Review* (2023-03-14). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVI2pW9UrhTD>
- [4] Ling Xiaoxiong, Wang Dingmin, Yuan Jian. A Calm Reflection on Technology Ethics and Academic Ethics After the ChatGPT Boom[J]. *Journal of Xinjiang Normal University (Philosophy and Social Sciences Edition)*, 2023, 44(04): 123-136.
- [5] Gao C A, Howardet F M, Markovet N S, et al. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers[EB/OL]. <https://doi.org/10.1101/2022.12.23.521610>
- [6] Li Jie. *Principles and Applications of Scientific Knowledge Mapping*[M]. Beijing: Higher Education Press, 2018.
- [7] R Core Team. R: A language and environment for statistical computing [computer software]. 2022, R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- [8] Van Eck N J, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping [J]. *Scientometrics*, 2010, 84(2): 523-38.
- [9] Xie Juan, Gong Kaile, Cheng Ying, et al. A Meta-Analysis of the Correlation Between Paper Downloads and Citations[J]. *Journal of the China Society for Scientific and Technical Information*, 2017, 36(12):
- [10] Qian Li, Liu Yi, Zhang Zhixiong, et al. Analysis of the Technical Foundation of ChatGPT[J/OL]. *Data Analysis and Knowledge Discovery* (2023-03-). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVI2pW9UrhTDCdPD66BsvSImaKhtD1BpWawVYd6jkT&uniplatform=NZKPT>
- [11] Guo Quanzhong, Zhang Jinyi. Technical Characteristics and Application Prospects of ChatGPT[J]. *China Media Technology*, 2023, (01): 159-160.
- [12] Zhu Guanghui, Wang Xiwen. Operational Mode, Key Technologies, and Future Vision of ChatGPT[J]. *Journal of Xinjiang Normal University (Philosophy and Social Sciences Edition)*, 2023, 44(04): 113-122.
- [13] Wu Xiaolin, Xing Yifei. Knowledge Replication or Innovation Stimulation? — Challenges and Opportunities of AI (ChatGPT) for Social Science Graduate

Education[J]. Journal of Guangxi Normal University (Philosophy and Social Sciences Edition) (2023-03-22).

[14] Zhong Binglin, Shang Junjie, Wang Jianhua, et al. Challenges of ChatGPT to Education (Roundtable Discussion)[J/OL]. Chongqing Higher Education Research (2023-03-02). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9qpqDiuWYnZMDDPAXo1QRcFhxOBE&uniplatform=NZKPT>

[15] Shen Chao. ChatGPT: Facilitating Higher Education Reform and Innovative Talent Cultivation[J]. Journal of National Academy of Education Administration, 2023, (03):

[16] Jiao Jianli. ChatGPT Boosts Digital Transformation of School Education — What to Learn and How to Teach in the AI Era[J]. Chinese Journal of Distance Education, 2023, 43(04): 16-23.

[17] Wang Youmei, Wang Dan, Liang Weiyi, et al. “Aladdin’s Lamp” or “Pandora’s Box”: The Potential and Risks of ChatGPT in Educational Applications[J]. Modern Distance Education Research, 2023, 35(02): 48-56.

[18] Zhang Xiaheng. Research on Digital Economy Development Based on New Generation AI Technology (ChatGPT)[J/OL]. Journal of Chang’an University (Social Science Edition) (2023-02-23). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2su>

[19] Zhou Zhibo. Introducing the ChatGPT Model into China’s Digital Government Construction: Functions, Risks, and Regulation[J/OL]. Journal of Shandong University (Philosophy and Social Sciences Edition) (2023-03-15). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTDCdPD66FeCiy>

[20] Wu Shiyou. Will ChatGPT Replace Social Workers?[J]. China Social Work, 2023, (06): 25-26.

[21] Liu Qin. Discussion on ChatGPT and Its Impact on Accounting Work[J]. Friends of Accounting, 2023, (06): 158-161.

[22] Zhang Zhixiong, Yu Gaihong, Liu Yi, et al. The Impact of ChatGPT on Library and Information Services[J/OL]. Data Analysis and Knowledge Discovery (2023-04-01). <https://kns.cnki.net/kcms2/article/abstract?v=3uoqIhG8C45S0n9fL2suRadTyEVl2pW9UrhTD>

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv — Machine translation. Verify with original.*