Visualization Analytics of Technology Empowered Education Research Using Citespace Software

Yuanyuan Fan*, Jianguo Li

School of Marxism, Huazhong University of Science and Technology, Wuhan, China

*fanyy@zzu.edu.cn, lijianguo@hust.edu.cn

Abstract. In recent years, with the rapid development of emerging technologies such as cloud computing, big data, artificial intelligence, and 3D computing, the research on technology empowered education has become a focus of attention for many interdisciplinary researchers and journals. This study used Citespace software and employed visualization and cluster analysis techniques to conduct a quantitative analysis of relevant literature on technology empowered education, summarizing the current status, progress, and trends of research on technology empowered education. In order to improve the quality of research, this study selected relevant literature on technology empowerment education research from Peking University Core and CSSCI included in the CNKI database as the sample objects. A total of 780 valid literature were screened, and bibliometric and CiteSpace knowledge graph analysis methods were used to visualize the number of publications, research authors, research organizations, and keywords, providing useful references for exploring the future development direction of technology empowered education research.

Keywords: Visualization Analytics; CiteSpace software; Mapping knowledge Domain; Technology empowered education research.

1 Introduction

The rapid development of information technology and the deepening of research on sustainable development theory have gradually made informatization and digitization the focus of global attention. Information science and information technology have widely and deeply influenced various fields of human society. Especially in the field of education, information technology has brought about significant changes in various aspects of education, from educational concepts to the educational process. As Abu-Salih B and Alotaibi S [1] argue, the mapping knowledge domain has established a dominant position in education as an extraordinary technological innovation. We must face the profound impact of modern information technology on education and fully tap into big data in education. With the development of information technology, research on technology empowering education must also keep pace. This is an important issue that researchers must urgently solve.

M. Yu et al. (eds.), *Proceedings of the 2024 5th International Conference on Big Data and Informatization Education (ICBDIE 2024)*, Advances in Intelligent Systems Research 182,
https://doi.org/10.2991/978-94-6463-417-4_44

Currently, the academic community is highly concerned about the research on information technology empowering education, with a high volume of publications. However, there is insufficient comprehensive research, making it difficult to form a scientific and systematic understanding, and even less quantitative research. To deepen relevant research, this article innovatively uses Citespace visualization software to quantitatively analyze the current status of research on information technology empowering education from the scientific perspective of knowledge graphs, grasp the research status of technology empowering education, and continuously promote the scientific and practical nature of research on information technology empowering education.

With the innovative development of advanced technologies such as artificial intelligence, big data, cloud computing, and computers, many advanced technologies and methods have shown significant advantages in data collection, analysis, and chart making in literature research. Among them, Cloud computing is developing into a firsthand archetype of large-scale distributed computing, which not only adds more power to internet technology [2], Mapping Knowledge Domain(MKD) combines similarity matrix calculation, clustering analysis, information visualization analysis techniques, etc., using visualized graphs to vividly display the current status, hotspots, and trends of related research. Citespace, as an information visualization software used for measuring and analyzing scientific literature data, displays the development trends of a certain discipline or knowledge field in a specific period of time, analyzes the evolution of cutting-edge fields, generates a systematic visualization knowledge graph, and reveals the development laws. Wang and Cai [3] conducted a visual analysis of the knowledge graph of Xi Jinping's ideological and political education ideology. Based on the advantages of MKD, Zhao [4] collected research papers on digital technology and sustainable development of the higher education industry both domestically and internationally from 2001 to 2022, conducted cluster analysis, and based on this, organized a research framework for digital talent cultivation in higher education, proposing transformation strategies to accelerate the future development of related fields. With the help of Citespace visualization analysis software, Chu and Wang [5] analyzed the research hotspots and future development trends of artificial intelligence empowering sports, and combined cluster analysis and literature analysis to identify five research hotspots: Healthy China, Integration of Sports and Medicine, Metaverse, Digital Technology, and Learning. This provides ideas and references for future related research and promotes the application and development of artificial intelligence in the field of sports. Xia and Wu [6] used Citespace visualization analysis software to conduct a visualization study on 281 relevant literature on "Artificial Intelligence & Ideological and Political Education" in the CNKI database of China National Knowledge Infrastructure (CNKI). They plotted the number of publications, keywords, and co-occurrence knowledge graphs, visualizing the transformation, risks, and paths of AI empowering ideological and political education in universities. Mamat M [7] employ Citespace to reveal research hotspots and frontiers of ATR, which provide insights into medical anatomy education. Trinh NTH [8] uses the VosViewer and Citespace tools to extract key insights into the role of higher education in national development in relation to human resources, economic growth, and other social issues. It can be seen that the deep integration of information technology and educational theory is beneficial for researchers to discover and solve problems. This study

can not only actively respond to new issues in technology empowered education research, but also help grasp the latest trends in technology empowered education research.

2 METHODOLOGY

Based on bibliometric methods, this article mainly uses the Citespace software developed by Dr. Chen Chaomei's team to draw a knowledge graph from the dimensions of authors, institutions, keywords, etc., making the data visual, dynamic, structured, and nodal. In the Time Slicing area of Citespace, set Years Perslice to 1 and perform time zone segmentation on the data to be analyzed, using 1 year as a slice to visualize the valid data. As there were no relevant literature published before 2011, the Cite Space time range was set to 2011-2023, with burst intensity as the default value, and node types selected as author, institution, and keyword. This study selected the most common threshold selection methods (TOP50, g-index (k=25), where g-index (k=25) is the default threshold in Citespace software, which refers to selecting analysis objects with a total of 252 citations from the first 25 papers) and drawing a knowledge graph.

This section introduces some concepts of clustering analysis and approaches that will be instrumental for Citespace software operation.

2.1 Cluster result analysis

The similarity visualization technology is well used in kind of literature analyses. As a popular statistics analysis tool, Citespace software is also operated based on the Cluster result analysis. The larger the word size represented by a node, the higher the frequency of occurrence, and the denser the connections between nodes, indicating a closer relationship between keywords. Q (Modularity Q) and the clustering silhouette index S (Silhouette) are the main judgment indicators for cluster result analysis. The Q value and S value are the main criteria for determining the clustering results of Cite Space. Q value greater than 0.3 indicates that the clustering is convincing, while an S value greater than 0.5 indicates that the clustering is reasonable [9].

Modularity is an evaluation index for network modularity, and its Q value calculation formula is defined as follows:

$$Q = \frac{1}{2m}\sum_{ij}(a_{ij} - p_{ij})\sigma(c_i, c_j) \tag{1}$$

where $A=a_{ij}$, is the adjacency matrix of the actual network; P_{ij} is the expected number of edges connecting node i and node j in the zero model; C_i and C_j represent the communities to which node i and node j belong in the network, respectively. When σ= At 1 o'clock, it indicates that i and j belong to the same club, otherwise σ= 0.

Silhouette is an indicator of network homogeneity. When S>0.5, it indicates reasonable clustering, and when S>0.7, it indicates reliable clustering. But if the S value is infinitely close to 1, it indicates that the homogeneity is too high, which is not conducive to analysis. The calculation formula for Silhouette is as follows:

$$s(i) = \frac{b(i)-a(i)}{max\{a(i),b(i)} \quad (2)$$

for one of the points i, calculate the intra cluster dissimilarity a (i): the average of the dissimilarity between the i vector and other points in the same cluster. Calculate the dissimilarity between clusters b (i): the minimum value of the dissimilarity between the i vector and other clusters.

$$a(i) < b(i), S(i) = 1 - \frac{a(i)}{b(i)} \quad (3)$$

$$a(i) = b(i), S(i) = 0 \quad (4)$$

$$a(i) > b(i), S(i) = \frac{b(i)}{a(i)} - 1 \quad (5)$$

If s(i) is close to 1, it indicates that the clustering of sample i is reasonable. If s(i) is close to -1, it indicates that sample i should be classified into another cluster. If s(i) is approximately 0, it indicates that sample i is on the boundary of two clusters.

2.2 Knowledge Domain Maps

The knowledge Domain Maps depends on the Euclidean distance. The distance of the cluster can be defined from the following formula:

$$E(X; D) = \sum_{i<j} d_{ij} ||\xi_i - \xi_j||^2 \quad (6)$$

In order to minimize the objective function, the following constraint should be considered:

$$\sum_{i<j} ||\xi_i - \xi_j|| = 1 \quad (7)$$

The average distance is denoted as

$$s = \frac{2}{n(n-1)} \sum_{i<j} ||\xi_i - \xi_j|| \quad (8)$$

where n is works of literatures.

The density of X (ξi , ξj) is formulated as

$$D(X) = \sum_{i<j}^{n} W_i k \left\{ \frac{||\xi_i - \xi_j||}{dh} \right\} \quad (9)$$

where W_i represents the frequency that the term i appears, k$\in$ [0, ∞) is function of the cluster kernel $k(t) = e^{-t2}$, h is a kernel parameter.

2.3 Data Source and Retrieved Content

This article uses Peking University Core and CSSCI journals indexed in the China National Knowledge Infrastructure (CNKI) database as search sources, and conducts a search on the theme of "Technology Empowered Education". As of November 29,

2023, a total of 1162 academic journals related to technology empowered education research were found. After literature review, a total of 780 valid articles were found after removing invalid ones.

3 DOCUMENT STATISTICAL ANALYSIS OF Technology Empowerment Education Research

3.1 Trends in annual publication statistics of literature

The interannual distribution and development changes of publication volume can reflect the development speed and research heat of the research topic within a certain period of time, which helps to grasp the annual distribution and change trend of the research topic as a whole [10].The earliest literature retrieved was a study by scholars Zhu Zhiting et al. [11] in 2011, as shown in Figure 1, especially after 2019, the research heat and literature volume in this field have accelerated, reaching 479 articles in 2023.

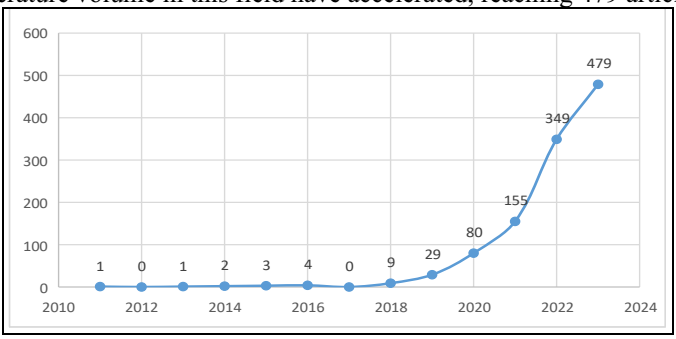

Fig. 1. Annual publication volume of literature

3.2 Quantitative Analysis of Research Authors

Analyzing the main authors and their collaborative relationships in technology empowerment education research can help us actively identify the hotspots in this research direction. The top five authors in the field of technology empowered education research are shown in Figure 2. According to statistics, from 2011 to 2023, the author with the highest number of publications in the research field was Zhu Zhiting (12 articles), followed by Xie Youru (7 articles) and Liu Bangqi (7 articles). They have achieved certain research results in the field of technology empowerment education.

According to the Price's Law formula ($M = 0.749 \times \sqrt{Nmax}$, where Nmax is the number of articles published by the author with the highest number of publications within the statistical period, and authors with more than M publications are referred to as core authors [12]). According to the 12 articles on technology empowering education with the highest number of publications by the first author, the following results can be obtained by applying the formula calculation:

$$M = 0.749 \times \sqrt{Nmax} = 0.794 \times \sqrt{12} = 2.584 \approx 3$$

According to the rounding principle, researchers with a publication volume of 3 or more articles in this field are the core authors, with a total of 21 people.

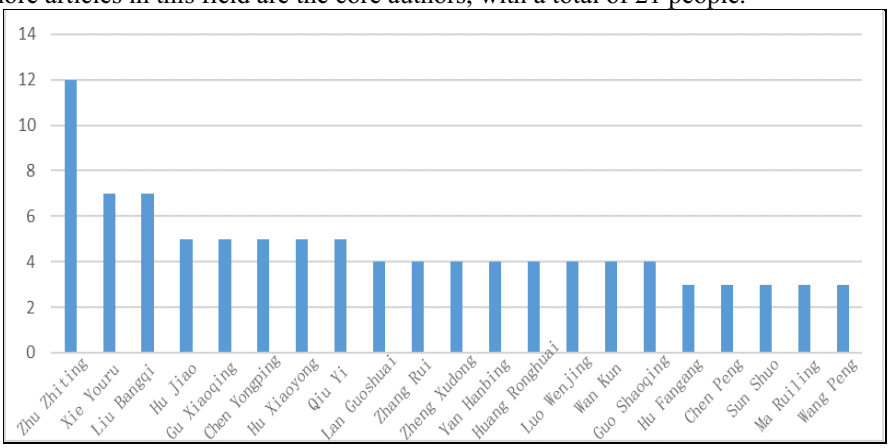

Fig. 2. The top five authors in the research field

As shown in Figure 3, the author co-occurrence graph network density presented by CiteSpace software is relatively low (0.0048), and the distribution among authors in this field is more dispersed, with less overall correlation among authors. Only a small number of scholars, including Zhu Zhiting and Hu Jiao, Xie Youru, Qiu Yi, Zhang Yue, Luo Wenjing, etc., have jointly explored issues related to technology empowerment education, and most researchers have not yet formed a stable cooperative relationship. At the same time, the authors have different focuses on the research field. Scholars such as Zhu Zhiting from the School of Open Education at East China Normal University focus on exploring the practical path and development direction of digital transformation; Xie Youru from the School of Education Information Technology at South China Normal University attach great importance to creating high-quality courses in universities in the context of the intelligent era; Overall, most research not only depends on the current historical context, but is also closely related to the research characteristics of different units.

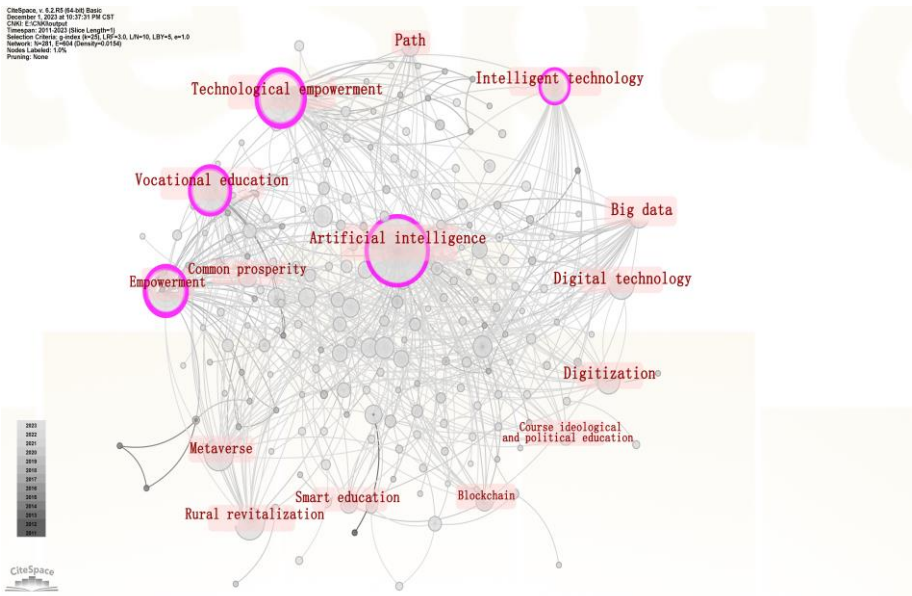

Fig. 3. The co-authorship network among productive authors

3.3 Quantitative Analysis of Organizations

According to the Price's Law formula, there are a total of 28 core institutions with a literature publication volume of ≥ 4. Through analysis of the main research institutions, the School of Education Information Technology at South China Normal University has the highest productivity within the discipline, publishing 18 articles. In addition, East China Normal University has formed important collaborations with universities such as Nanjing University of Posts and Telecommunications and Jiangsu Normal University in this research field. Meanwhile, the cooperation between different departments of major universities is relatively close. Centrality is an indicator in CiteSpace that measures the importance and influence of literature, with values ranging from 0 to 1. The higher the value, the more important the node is.

As shown in Figure 4, the research institution with the highest centrality is the Open Education College of East China Normal University, with a centrality of 0.01. The co-occurrence graph network has a low density (0.0044), and cooperation between different institutional groups is not close, with only small-scale cooperation between institutions. Moreover, current research on technology empowered education mainly focuses on teacher training institutions, with education colleges or educational information technology colleges as the main focus. This is mainly because the research subject has a high degree of compatibility with teacher training institutions and education majors.

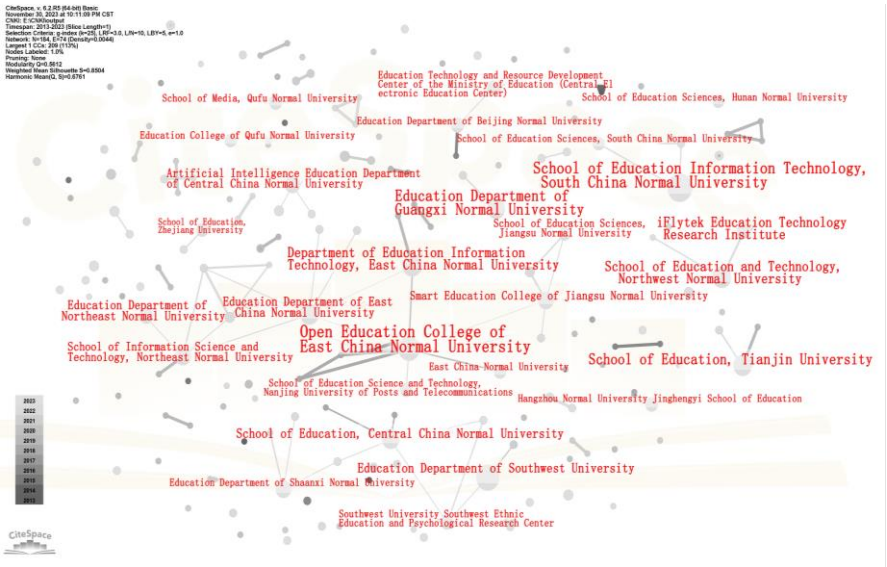

Fig. 4. Collaboration network among main research organizations

3.4 Keyword co-occurrence analysis

The co-occurrence map of keywords in technology empowered education research is shown in Figure 5. In the visualization map, there are 280 nodes and 603 connections, with a network density of 0.0154, indicating a weak degree of connection in keyword research. Keywords with a centrality greater than 0.1 are considered core keywords. As shown in Figure 5, artificial intelligence (0.36) has the highest centrality among the keywords, followed by technology empowerment (0.26), empowerment (0.24), vocational education (0.19), and intelligent technology (0.12).

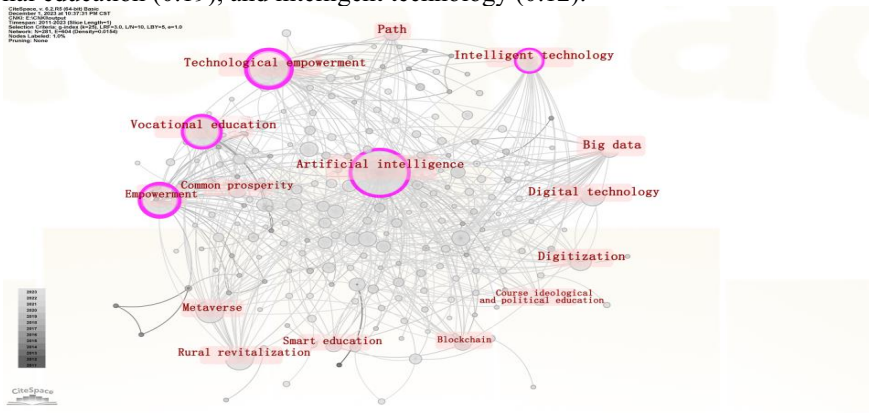

Fig. 5. Keyword co-occurrence network

3.5 Keyword clustering analysis

Cluster analysis is the process of classifying data without classification information based on similarity using certain methods [13]. Keyword clustering can reflect the hot topics in the field, and through these clustering, the development laws and new directions of this discipline can be seen [14]. Q and S values are the main judgment indicators of CiteSpace for clustering results. In this study, Q=0.561>0.3 indicates that the clustering is convincing, and S=0.850>0.5 indicates that the clustering is reasonable. Using Citespace software to draw a keyword map of education research literature empowered by technology, as shown in Figure 6, a cluster of # 0 artificial intelligence, # 1 intelligent technology, # 2 vocational education, # 3 technology empowerment, # 4 digitization, # 5 empowerment, # 6 curriculum ideology, # 7 education equity, etc. has been formed. Research on technology empowerment education revolves around these clusters.

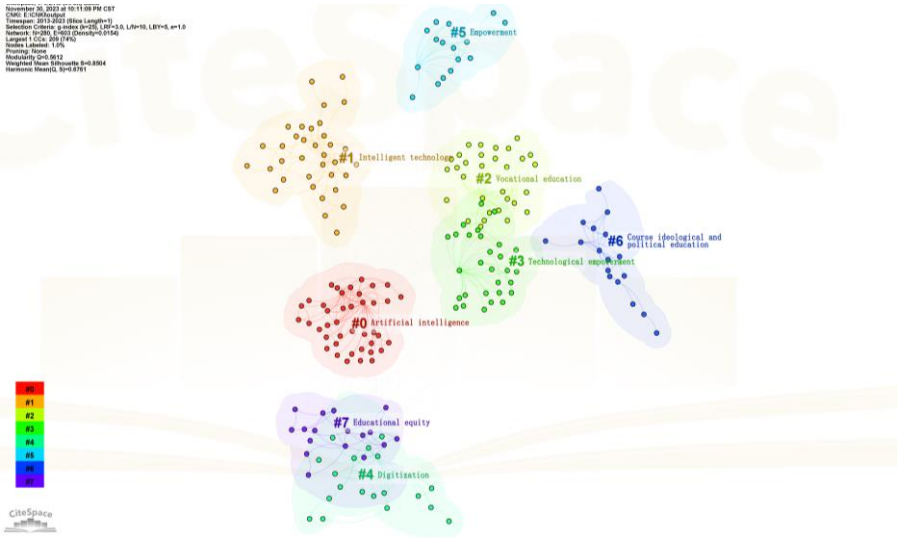

Fig. 6. Keyword clustering diagram

3.6 Keyword mutation analysis

As shown in Figure 7, in the keyword emergence chart, Year represents the year when the keyword first appeared, Begin and End represent the starting and ending years when the keyword was used as a frontier, and Strength represents the intensity of the emergence, reflecting the research frontier at different time periods. The research in this field from 2011 to 2023 can be mainly divided into two stages. The first stage is from 2011 to 2021, which is a steady growth stage of research and lays the foundation for subsequent research. The second stage is from 2021 to 2023, with the main research focus on teacher, vocational education, smart classroom, and educational transinformation.

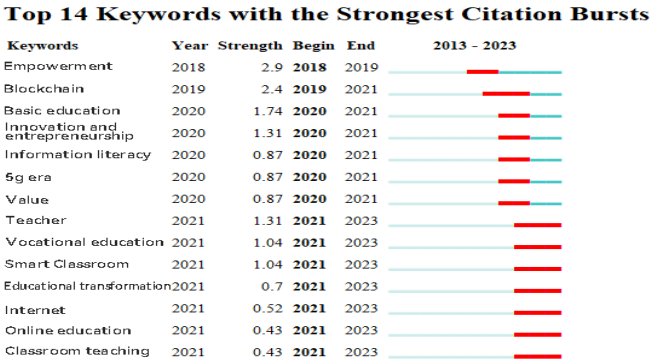

Fig. 7. Keyword emergence chart

4 Conclusions

This article uses Citespace software to analyze the research on technology empowered education, including annual publication volume, authors, institutions, keyword clustering and keyword highlighting. The following conclusions are drawn:

(1)By analyzing the number of publications from 2011 to 2023, it can be seen that current research on technology empowered education is characterized by fragmentation and fragmentation. The core authors are represented by scholar Zhu Zhiting, and most scholars are still in a state of "fighting alone". However, the cooperation between authors needs to be strengthened, and collaborative scientific research has not yet formed. From the analysis of the institutions published in the article, it can be seen that normal universities have an absolute advantage, but research institutions are relatively scattered, and there is less cooperation between institutions. Future research aims to cultivate interdisciplinary perspectives, particularly by strengthening communication and exchange with professionals such as information technology, artificial intelligence, and big data. By fully utilizing research databases and other advantageous platforms, we actively seek cooperation and strive to form an academic community.

(2)From the perspective of research methods, the research tends to be theoretical and lacks certain empirical support. In recent years, with the development of information technology, big data, and artificial intelligence, research hotspots such as metaverse, science and education integration, ChatGPT, virtual simulation, VR technology, and smart platforms have gradually become areas that people attach importance to in exploring the path of becoming a technological and educational powerhouse. However, existing research mainly focuses on concepts, characteristics, influencing factors, lacking micro and empirical research on technology empowered education. Future research should break the inherent research framework and conduct quantitative analysis of technology empowered education. Using technologies such as big data and cloud computing to accurately analyze the level of information literacy of teachers, using information methods to fully explore big data in various levels of education, accurately grasp the real needs

of students for education, and empower the promotion of high-quality development of education.

(3)From the perspective of research hotspots, the research content is slightly scattered, and the research direction needs to be further deepened. According to the keyword clustering graph, it can be found that current technology empowered education is more focused on vocational education and curriculum ideology. In fact, from primary and secondary schools to universities, from classroom teaching to extracurricular practice, from teaching reform to educational reform, from basic education to vocational education, and then to higher education, from physical education to labor education, education always requires effective technological empowerment. Future research should continuously deepen the research content and objects of technology empowerment education, broaden horizons, and strengthen research from multiple perspectives and dimensions.

Acknowledgement

This work is supported by National Social Foundation of China under Grant 20BKS103, Research on the Discourse Power and Discourse Power Construction of Ideological and Political Education for Post-2000 College Students.

References

1. Abu-Salih B and Alotaibi S. “A systematic literature review of knowledge graph construction and application in education”. Heliyon, vol.1, no.10, 2024.
2. John Kwao Dawson, Ben Beklisi Kwame Ayawli, Sylvester Agyemang, Philemon Baah, and Samuel Akyeramfo-Sam, “Ensuring Cloud Data Security Using the Soldier Ant Algorithm”, Journal of Advances in Information Technology, vol.14, No.1, February 2023.
3. Shujun Wang and Xiaoshen Cai, “Literature analysis of Xi Jinpings important expositions on ideological and political education based on cluster analysis”, 2021 International Conference on Education Knowledge and Information Management, pp:792-796, 2021.
4. Yixin Zhao, “Research and Transformation Direction of Digital Empowerment of Higher Education Talent Training - Visual Analysis Based on Knowledge Graph (2001-2022)”, East China Science & Technology, vol.03, pp:146-148, 2023.
5. Jiaze Chu and Yong Wang, “Research hotspots and trends in artificial intelligence empowering sports based on CiteSpace visualization analysis”, Science & Technology of Stationery & Sporting Goods, vol.14, pp:146-148, 2023.
6. Mengxuan Xia, Chunyang Wu and Aiqin Lin, “Review and Prospect on the Research of Ideological and Political Education in Colleges and Universities Empowered by Artificial Intelligence—Data Analysis Based on CiteSpace Knowledge Graphs”, Journal of Luohe Vocational and Technical College, vol.22, no.03, pp:146-148, 2023.
7. Mamat M, Li L,Kang S, Chen Y. “Emerging trends on the anatomy teaching reforms in the last 10 years: Based on VOSviewer and CiteSpace”, Anat Sci Educ, vol.13, 2024.
8. Trinh NTH. “Higher Education and Its Role for National Development. A Research Agenda with Bibliometric Analysis”, Interchange, vol.54, pp: 125-143, 2023.
9. Sup Xujian Sup, Sup Dai Fangfang Sup, Sup Pan Wenxue Sup. “Visual Analysis of Hotspots and Cutting-edge Trends of Community TCM Service Research in China in the Context of Healthy China”, Chinese General Practice, vol.34, pp:4343-4350, 2023.

10. Jinjuan Duan and Aiqi Xuan. “Research Trends of Intangible Heritage Industrialization Protection Based on Knowledge Graph”. Packaging Engineering, vol.42, no.6, pp:162-170, 2021.
11. Zhiting Zhu and Bin He, “Analysis of the National Education Technology Plan 2010 in the United States”, China Educational Technology, vol.06, pp:16-21, 2011.
12. Sup Kangmin Sup, Sup Wu Congli Sup and Sup Wu Benqing Sup, “Research hotspots and visualization analysis in hospital statistical information data analysis in the era of big data”, China Health Statistics, vol.02, pp:219-223, 2023.
13. Sup Yang, Lulu Sup, Sup Ma Hengyun Sup and Sup Yu Song Sup, “The Development and Prospects of Rural Homestead Research in China in the Past 20 Years * - Visual Econometric Analysis Based on CNKI and CiteSpace”, China Agricultural Resources and Regionalization, vol.06, pp:181-192, 2023.
14. Hualing Miao, Cheng Guo, Mingchen Wei, “The Characteristics and Prospects of Research on the Psychological Quality of Children and Adolescents: Visual Analysis Based on CiteSpace”, Journal of Southwest University (Social Sciences Edition), vol.04, pp:206-217, 2023.