



# Analysis of the development trend of teaching application research of big data technology in recent 10 years

Huiyuan Luo

College of Education Science, Bohai University, Jinzhou 121013, China

E-mail: luohuiyuan98@qq.com

**Abstract.** With the deepening of the application of information technology in the field of education and teaching, a large number of educational data can be produced. Based on 611 journal papers related to the teaching application of big data technology collected in CNKI database, this paper makes an in-depth discussion on the hot spots and development trends of the teaching application of big data technology in China in the past ten years by using the information visualization software CiteSpace and the analysis methods of spatiotemporal knowledge graph and content knowledge graph. The analysis finds that the characteristics of big data technology in teaching application research are as follows: Since 2013, the application of big data technology in the field of education has achieved remarkable results, and entered a stage of large-scale and rapid development during 2014-2017. The research content mainly focuses on big data, application, learning analysis, data mining, data analysis and smart education, etc., which has become a research hotspot. The application of big data technology in the field of education mainly involves the aspects of learning analysis, personalized education, teaching evaluation and student management. In addition, factors such as educational policies, educational resources and teachers' professionalism also have an impact on the application of big data technology in the field of education.

**Keywords:** Big data technology; Teaching application; CiteSpace

## 1 Introduction

With the continuous advancement of information technology, the rate of data generation, processing and use in human society has increased alarmingly. Through the exchange, integration and analysis of data, the scope of human's recording, measurement and understanding of the objective world is also constantly expanding. These massive amounts of data, known as "big data", can be used to mine new knowledge, discover new rules and create new value. <sup>[1]</sup>The emergence and rapid development of big data technology has had an important impact on all walks of life, which cannot be ignored. In the field of education, big data technology has also set off a new technological revolution, and put forward the concept of "big data for education". Educational big data refers to the data set with great application potential that is generated by traditional

© The Author(s) 2024

G. Guan et al. (eds.), *Proceedings of the 2023 3rd International Conference on Education, Information Management and Service Science (EIMSS 2023)*, Atlantis Highlights in Computer Sciences 16, [https://doi.org/10.2991/978-94-6463-264-4\\_18](https://doi.org/10.2991/978-94-6463-264-4_18)

collection methods, portable electronic products and application software and other integrated systems in the whole process of education and teaching. <sup>[2]</sup>Educational big data can accurately analyze and study students' academic performance, exam results and future career development and planning. The US government invested 200 million US dollars in 2012 to carry out large-scale data reform of the education system, and strive to better study local education methods through big data technology, so as to establish a more sound and effective education platform.

There is no doubt that big data technology will bring profound influence and comprehensive reform to education: the first is the reshaping of teacher-student relationship. From teachers' "teaching" as the core, to a new teaching relationship with teachers and students' joint in-depth development as the core. Secondly, make personalized education possible. The big data teaching platform can collect students' data in detail, recommend the best learning materials and the most effective learning plans for different students, and tailor learning plans for each student. Finally, it will bring about profound changes in all aspects of educational activities. For example, the focus of teaching has shifted from offline to online platforms, students can choose the courses they are interested in independently, and enjoy personalized and high-quality resources.

In recent years, various emerging technologies continue to emerge, such as meta-universe technology, generative artificial intelligence technology led by chatGPT, etc. Their birth and application are mostly closely related to big data technology, and the education and teaching application of big data technology has received further attention. Therefore, this paper uses the scientific knowledge graph analysis method to carry out multiple researches on the domestic literature on the teaching application of big data technology retrieved in the past decade, that is, CiteSpace knowledge graph is used to carry out visual analysis on the teaching application research of big data technology, and explore the development trend of the application research of big data technology in the field of education and teaching. In order to effectively help promote the deep application of big data technology in education and teaching.<sup>[3]</sup>

## 2 Research Program

### 2.1 Research Tools

CiteSpace is a visual analysis tool based on literature data, developed by a research team led by Professor Chaomei Chen at Drexel University. The tool can help researchers discover and analyze information about hot areas of academic research, key authors, citation networks and so on in large-scale literature data. CiteSpace helps researchers better understand the development trends and knowledge structure in the academic field by visualizing the relationships between literature. The tool has been widely used in academia and scientific research institutions, and provides strong support and help to researchers. The CiteSpace software provides several function buttons, including Keyword, Cited Journal, Cited Author and Cited Reference, among others. <sup>[3]</sup>These function buttons can help users conduct targeted search and analysis in large-scale literature data. Through the keyword function, users can enter keywords to find literature related to a specific topic; The journal citation function can show how literature published in a

journal has been cited by other literatures, helping users understand the influence and research hotspots of the journal; The author citation function can show the situation of a certain author's literature being cited by other literatures, helping users to understand the academic influence of the author; The literature citation function can show the situation of a certain literature being cited by other literatures, helping users to understand the influence of the literature and the citation network. These function buttons make it more convenient for users to carry out literature analysis and knowledge discovery.

The specific operation steps of CiteSpace analysis of big data technology teaching application research are as follows: Use the data format conversion tool of CiteSpace to convert the documents exported in CNKI in Refworks format to the data format recognized by CiteSpace;<sup>[3]</sup>The time span is set to 2013 to 2023, with an interval of 1 year; Choose a cut connection method for Pathfinder to simplify the network structure and highlight important features; Finally, the final analysis map is presented in the form of Cluster View-static and Show Merged Network visualizations.<sup>[4]</sup>

## 2.2 Data Sources

The research focuses on word frequency, clustering, hot spots and emergent words analysis of the key words of domestic big data technology teaching application research. Before the analysis, first of all, the key words in the original literature data should be unified screening and analysis.<sup>[3]</sup>In data collection, this study will use CNKI database as the main literature source. CNKI, short for China National Knowledge Network, is one of the largest academic literature databases in China, covering a large number of literature resources in various subject areas. The reasons for choosing CNKI as a search database are as follows: 1. Abundant literature: CNKI has collected a large number of academic journals, academic dissertations, conference papers and other literature resources, which can provide comprehensive research literature support;<sup>[5]</sup>2. Wide coverage: CNKI covers literature resources in various disciplines, including education, computer science, data science and other related fields, which can meet the research needs of this study on the teaching application of big data technology in China; Powerful retrieval function: CNKI provides advanced retrieval function, which can conduct accurate literature retrieval according to specific search conditions, which is helpful to screen out the literature related to this research. By using CNKI database for data collection, this study can obtain relatively comprehensive and reliable domestic literature on the teaching application of big data technology, which provides a reliable data basis for the subsequent analysis of word frequency, clustering, hot spots and emergent words.

The data collection method for this study is as follows: Select the type of "advanced search" and "subject" search in CNKI database, and the search condition is "Education application & Big Data"<sup>[3]</sup>. By July 7, 2023, a total of 878 relevant literatures have been retrieved. After manual screening, eliminating reports, conference notices, documents, calls for papers, article titles and other documents inconsistent with the content of this study, 611 valid documents were obtained.

### 3 Space-time knowledge graph and its analysis

#### 3.1 Time distribution graph of teaching application research of big data technology

In order to investigate the research achievements of big data technology in the field of education application, the literature published from 2013 to 2023 was counted in this study, as shown in Figure 1. Figure 1 shows that from 2013 to 2021, the research literature on the teaching application of big data technology shows an overall rising trend, and the number of published papers increases sharply from 2020 to 2021, reaching a peak in 2021. On the one hand, the advent of the era of big data makes the collection, analysis and utilization of educational data more convenient and efficient, which is more conducive to the continuous development of big data technology in the field of teaching application; On the other hand, 2013 is the "first year of big data" in China. In this year, a wave based on big data is emerging in the education sector to promote education reform and innovative development. The relevant research of big data technology in teaching has developed rapidly and achieved remarkable results.<sup>[3]</sup>From 2020 to 2021, the world is ushered in the era of epidemic, and online teaching has become the main teaching method in the daily life of teachers and students. Online education based on the Internet, media and information technology makes it possible to collect the whole process of big data in education, and catalyzes the rapid development of big data technology in the field of teaching application.

Figure 1 also shows that the research on the teaching application of big data technology in China shows a downward trend from 2021 to 2023, and the number of relevant literature publications in 2022 returns to a level close to that of 2017-2020, and the number of publications in the first half of 2023 is less than half of the average number of the past seven years. This is mainly because big data technology was born earlier, and after more than a decade of continuous discussion and research by domestic and foreign scholars, it has gradually matured. In recent years, with the continuous emergence of a variety of emerging technologies, research hotspots have shifted, and big data technology is usually directly applied as the basis for research on a variety of emerging technologies, and is rarely studied and discussed separately.

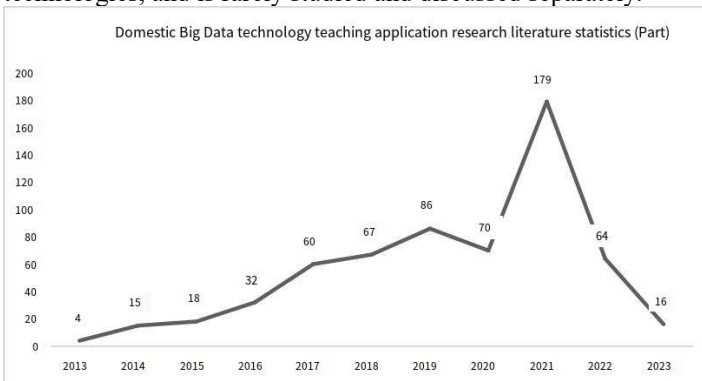
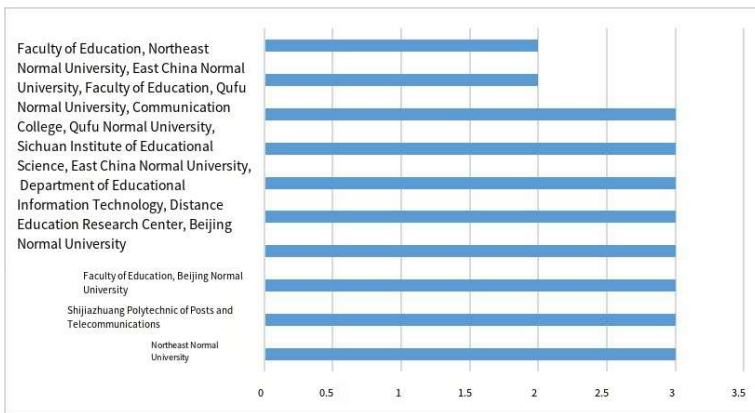


Fig. 1. Literature statistics on the teaching application of big data technology in China

### 3.2 Spatial distribution map of applied research on teaching of big data technology

In order to study the core academic groups and institutions of big data technology in teaching application in China in the past decade, this study made statistics on the papers published by various research units in this field, among which the situation of high-yield institutions is shown in Figure 2. [6]As shown in Figure 2, Northeast Normal University, Shijiazhuang Polytechnic of Posts and Telecommunications, Faculty of Education of Beijing Normal University, Distance Education Research Center of Beijing Normal University, Department of Educational Information Technology of East China Normal University, Sichuan Institute of Education Sciences, Communication College of Qufu Normal University, and Faculty of Education of East China Normal University all occupy the top 8 with 3 publications. It shows that these 8 research institutions have a great space for development in the teaching application of big data technology. In addition, the number of published papers of the top 10 institutions is not different, which indicates that the application research of big data technology in the field of education has been attached great importance by many departments and institutions.



**Fig. 2.** High production institutions of big data technology in teaching application research

In order to study the cooperation between different institutions, graduate students have become the cooperation map of institutions for teaching and applied research of big data technology, as shown in Figure 3. [3]In the graph, each node represents the name of an institution, the area of the node represents the amount of papers issued by the institution, the color of the node represents the time of publication, the size of the mark font represents the centrality of the institution, and the edge represents the cooperative relationship between the institutions. [7]There are 407 nodes and 107 connections in the institutional cooperation network, and the overall density of the network is only 0.0013, indicating that the research groups of teaching and application research of big data technology in China are scattered, and the authors of different institutions lack cooperation, so there is no more influential research group. [3]

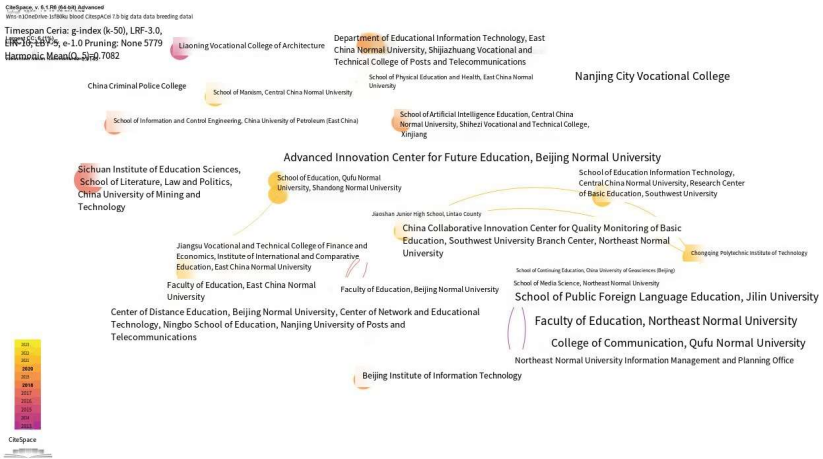


Fig. 3. Cooperation map of teaching and application research institutions of big data technology

In order to study the research status of educational big data in different institutions at different times, the graduate students formed the institutional research time series Atlas. The time series in the graph are displayed in different colors, and each color corresponds to the node age ring color (see Figure 4).<sup>[3]</sup>Figure 4 shows some high-yield institutions, among which Northeast Normal University first began to participate in the teaching application research of big data technology in 2013 and continues to do so until now, indicating that the institution has always maintained a certain degree of attention and strong stability in the field of teaching application research of big data technology. Beijing Normal University, East China Normal University, Sichuan Institute of Education Science, Qufu Normal University, Shijiazhuang Vocational and Technical College of Posts and Telecommunications and other institutions began to pay attention to the research in the field of teaching application of big data technology in 2016, and the research groups have been relatively stable since then. This indicates that the research on the teaching application of big data technology in China ushered in a boom from 2013 to 2023, and more high-yield institutions carried out related research in this field during this period.

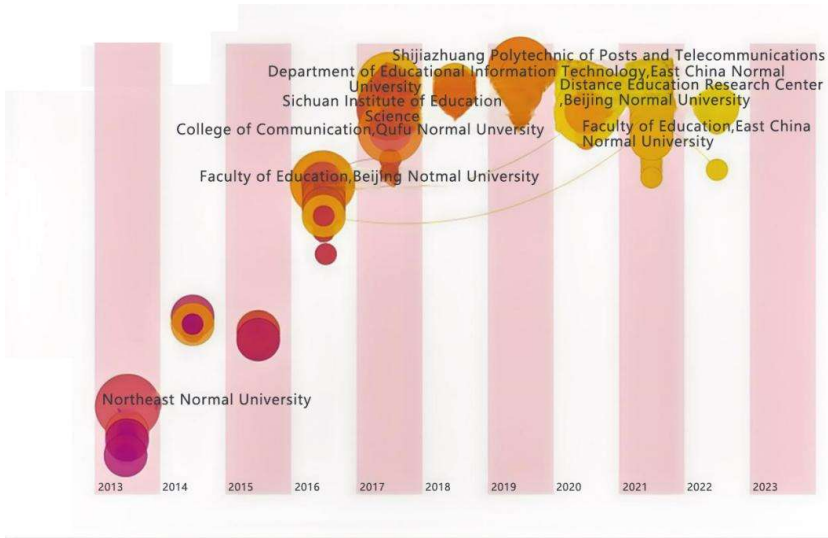


Fig. 4. Time series of research institutions on teaching and application of big data technology

## 4 Content Knowledge graph and its analysis

### 4.1 Analysis of hot topics in teaching application research of big data technology

In theory, centrality is an index to measure the importance of nodes in the network, which reflects the influence or power of nodes on the network. [8]In a keyword co-occurrence network, if a node has a higher centrality, it means that the node has a higher importance in the field. This usually means that the node has a higher degree of correlation with other nodes, that is, the keyword co-occurrence frequency is higher. Therefore, the degree of centrality can be used to evaluate the importance of a node in a particular domain. As shown in Table 1, the keywords with high frequency in the research literature on teaching application of big data technology in China in the past decade include "big data", "application", "university", "education management", "data mining", "informatization", "education", "higher education", "learning analysis", "university education", etc. It can be seen that there are more researches on big data technology in the field of higher education teaching application, but fewer researches in primary and secondary schools and preschool education.

Table 1. Keyword co-occurrence frequency, centrality and age (part) (span: 1 year)

Serial Number	Frequency	Centrality	Vintage	Keywords	Serial number	Frequency	Centrality	Vintage	Keywords
1	349	1.25	2013	Big data	17	12	0.01	2014	Educational apps

2	51	0.05	2014	Apply	18	11	0.04	2014	Data analysis
3	38	0.01	2015	Colleges and universities	19	10	0.01	2021	Precision teaching
4	27	0.02	2016	Education Administration	20	10	0.02	2016	Teaching model
5	24	0.04	2013	Data Mining	21	10	0.04	2017	Smart Education
6	22	0.01	2016	Informationization	22	9	0	2016	Applied Research
7	20	0	2016	Education	23	9	0	2014	Education reform
8	20	0.01	2016	Higher education	24	9	0	2016	Strategies
9	19	0.05	2013	Learning analysis	25	8	0	2014	Cloud Computing
10	16	0.02	2014	Higher education	26	8	0.01	2017	College students
11	15	0.02	2015	Vocational education	27	8	0.03	2016	Talent development
12	15	0.01	2014	online education	28	7	0.01	2017	Innovation
13	12	0.02	2017	Internet	29	7	0.01	2017	Educational technology
14	12	0.03	2016	Teaching reform	30	7	0.03	2019	Information Technology
15	12	0.03	2016	Artificial Intelligence	31	6	0	2014	Research
16	12	0.01	2016	Higher vocational colleges	32	6	0	2014	Education field

CiteSpace's keyword clustering feature allows for a better understanding of hot topics and trends in a field of study. <sup>[9]</sup>By analyzing the co-occurrence relationship of keywords in literature, the function clusters relevant keywords together to form nodes, and displays the size, color and thickness of nodes according to the co-occurrence frequency and occurrence time of keywords. In the generated keyword cluster graph, the circle represents the keyword node, and the size of the circle reflects the discussion frequency of the corresponding topic. If the circle area of a topic is large, it indicates that the topic has been widely discussed and concerned in the research field; The color and thickness of the node rings represent the time when the keywords appeared, the color of the keyword nodes that appeared earlier is lighter, the color of the keyword nodes that appeared



later is darker, and the thickness of the node rings also increases over time. The literature data of teaching and application research of big data technology derived from CNKI were processed, and the age was divided into 1 year to obtain the clustering map of the literature of teaching and application research of big data technology.<sup>[3]</sup>

Through the analysis of the clustering results (as shown in Figure 5), "big data" has been at the center of China's research in the past 10 years, which indicates the importance and foundation of big data technology; Followed by "higher education", "universities", and then "data analysis", "education", "data mining", "teaching mode", "education data", "information technology", "learning analysis", "deep learning", etc., basically consistent with the keyword co-occurrence frequency statistics. The higher the centrality,<sup>[3]</sup> the more important the node. In the research nodes of big data technology teaching application research in China in the past decade, the higher are "big data" (1.25), "application" (0.05), "learning analysis" (0.05), "data mining" (0.04), "data analysis" (0.04), "smart education" (0.04) and so on.

Figure 5 shows that there are a total of 482 nodes and 1030 connections in the application research of big data technology education in China, and the overall network density is 0.0089. On the whole, the keyword co-occurrence network is relatively loose and the density is low. <sup>[10]</sup>It is suggested to further promote the academic collaboration of relevant researchers in the future, and enhance the concentration of their research topics to conduct more accurate and in-depth research.

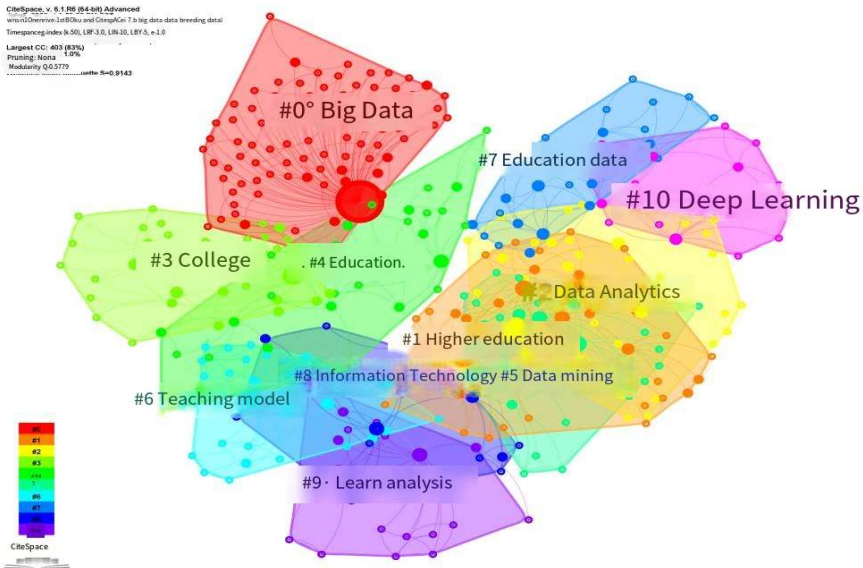


Fig. 5. Clustering map of keyword co-occurrence frequency in big data technology

#### 4.2 Research trend of big data technology teaching application research

Abrupt words refer to words that have been widely discussed or used with a high frequency in a short period of time. By observing the change in word frequency of mutant words, it is possible to judge the frontiers and trends in the field of research.<sup>[3]</sup>Based on the results of CiteSpace correlation analysis, the abrupt topic, prominence rate and cited history curve in the teaching application field of big data technology are obtained, as shown in Figure 6. "Learning analysis", "data mining" and "online education" are the research hotspots in the field of teaching application of big data technology, among which "learning analysis" is mainly reflected in 2013-2016, "data mining" is mainly reflected in 2013-2015, and "online education" is mainly reflected in 2014-2017. This shows, to a certain extent, that the current research frontier of the teaching application research of big data technology in China is mainly reflected in the field of learning analysis, data mining and online education.

Top 25 Keywords with the Strongest Citation Bursts



Fig. 6. Emergence rate of keywords in teaching application research of big data technology

Based on the analysis of cluster graph, the time series statistics of the leading keywords in the teaching application research of big data technology are conducted according to time segments, and the results are shown in Figure 7. The research on the teaching application of big data technology in China can be roughly divided into three stages<sup>[3]</sup>: the first stage is the initial stage (2013-2014), which mainly involves the application of big data technology, learning analysis and data mining; The second stage is the development period (2014-2016), which involves the application of big data technology in various fields of education, such as online education, education reform, education management and higher education. The third stage is the maturity stage (2017-2023), during which there are fewer new topics and more research topics that have been proposed in the past are further discussed. At the same time, this stage is closely related to the progress and development of the Internet and information technology, and the research focus is on the related applications combined with emerging technologies.

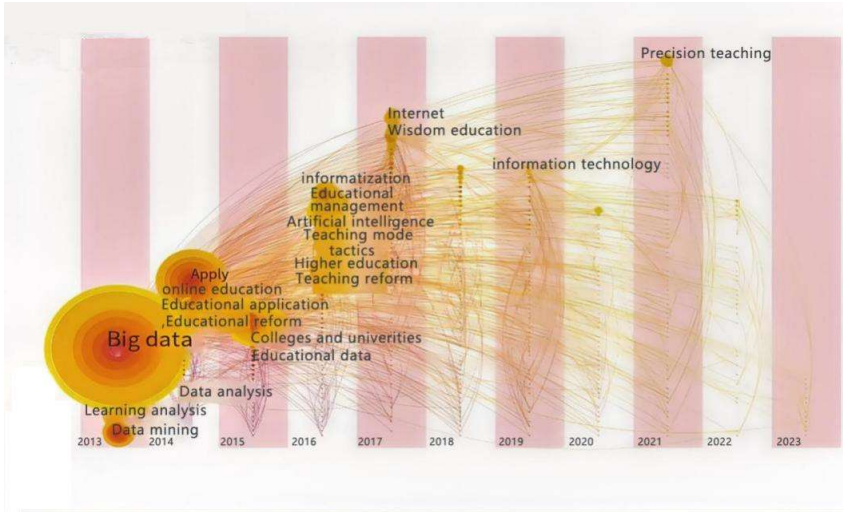


Fig. 7. Sequence map of key words in teaching application research of big data technology

## 5 Conclusion and thinking

### 5.1 Research Conclusion

Using CiteSpace software, this paper carries out multi-level analysis and visualization research on the knowledge graph generated by the research literature on the teaching application of big data technology in CNKI database from 2013 to 2023 and the related data, and finally draws the following conclusions:

(1) The time distribution graph shows that: the research on the teaching application of big data technology has extremely important value. The earliest research can be traced back to 2008, but concentrated research began to appear in 2013. From 2014 to 2017, the field entered a stage of large-scale and rapid development, with a significant improvement in the quantity and quality of studies. During this period, research on teaching application of big data technology has gained wide attention and recognition. Scholars began to explore how big data technology could be applied in the field of education to improve teaching effectiveness and student learning outcomes. The research focuses on the application of big data analysis, data mining, machine learning and other technologies in teaching, as well as how big data can be used to personalize teaching, assess student performance and improve teaching strategies.

(2) The spatial distribution map shows that there are more institutions participating in the teaching and application research of big data technology, and the scientific research team is growing, but the cooperation between the institutions is less, and the research institutions and leading scholars with sufficient influence are lacking. In the future development, it is necessary to strengthen the cooperation between institutions and scholars, establish a long-term and effective guarantee mechanism, so as to keep the attention of scientific research institutions and scholars on research, and eventually

form the mainstream research direction, and promote the virtuous circle of big data technology in teaching application research.

The keyword co-occurrence graph shows that the hot spots of teaching application research of big data technology are: "big data", "application", "learning analysis", "data mining", "data analysis", "smart education" and so on. The structure of keyword co-occurrence network is relatively loose and the density is not high, so researchers need to strengthen the accurate and in-depth research on a certain topic in the future.

(3) The time-series map of research frontiers shows that the frontiers of domestic research on teaching application of big data technology are reflected in the fields of "learning analysis", "data mining" and "online education"; From the early learning analysis and data mining, to the application of big data technology in teaching management, education reform, teaching mode and other fields, and then to the combination of big data technology with the Internet and information technology, the development of big data technology teaching application research also reflects the continuous integration and interaction of advanced technologies such as data mining and information technology. The time series map also reflects that: the research topic of big data technology teaching application research is relatively single, and the research field and scope are small, which can enhance cross-section, interdisciplinary and cross-field cooperation and exchange, as well as the integration and application of artificial intelligence-related technologies, and open up new research paths.

## 5.2 Research Thinking

The application of big data technology has opened a new era of the development of education and teaching practice, and promoted many changes in the field of education. However, with the deepening of the application, there are still many problems worth discussing and thinking about: First, how to realize the sharing and integration of cross-field education database is one of the urgent problems in education practice. Based on the massive data sources, the application of big data technology should establish a database containing a variety of different types of information, the mere use of education-related databases is not enough, and can not fully reflect the various information related to the object of education analysis. Second, education big data practitioners must have a certain amount of data literacy in order to effectively conduct big data analysis. First, practitioners can obtain a variety of high-quality teaching resources, and effectively manage and store them; Secondly, practitioners should have good data analysis ability, with certain data statistics and analysis ability; Finally, the data information can be interpreted, and the use of model analysis of the data for depth mining. Although there are still many problems to be solved, educational decision-making guided by educational big data has begun to become an important basic guarantee for modern education and teaching reform, and has become an unchangeable development trend of educational information. In order to make the cause of education big data develop rapidly, it is necessary to systematically integrate the whole society and educational resources, and it is more urgent to enhance the quality of big data of front-line education personnel.

At present, the research on the teaching application of big data technology is usually aimed at higher education. The reason is that face-to-face classroom teaching can not

fully meet the development needs of students' personal learning career and career path in the field of higher education and adult education. Independent learning based on online education has become one of the main learning methods in the daily life of most students. The wide application of online education greatly facilitates the collection and extraction of big data in education, making big data technology more widely used in higher education. With the continuous improvement of the level of social science and technology, as well as the significant impact of the epidemic era on the field of education, online education has gradually been accepted by a wider group of students, bringing new opportunities to the application of big data technology in primary and secondary education and even preschool education.

## References

1. Zhang Yannan, Zhao Zhongjian. Ethical thinking on the application of big data in education [J]. *Global Education Outlook*, 2016, 45(01): 48-55+104.
2. Xu Lei, Nie Fengying. Research status and technology Path of educational big Data application in educational management [J]. *Higher education in Jiangsu*, in 2020, No. 238 (12) : 69-73. The DOI: 10.13236 / j.carol carroll nki jshe. 2020.12.013.
3. wang Juan, shi-chao Chen, Wang Linli, etc. Research hotspot and trend analysis of Big Data in Education based on CiteSpace [J]. *Modern Educational Technology*, 2016, 26(02): 5-13.
4. Duan Aifeng, Wu Yuqing. Knowledge Graph Analysis of computational Thinking research in China -- Visualization Analysis based on CiteSpace [J]. *Journal of hebei university adult education college*, 2019, 21 (01) : 47-55, DOI: 10.13983 / j.carol carroll nki jaechu. 2019.01.007.
5. Liu Chun, Liu Hong, Wang Yanan. [8] Research on Application strategy of logical search expression for professional search [J]. *Information Exploration*, 2015(06): 70-73+135.]
6. Wang Tian, Yuan Mei. Hot spots and context evolution of residential education research in China since the 21st century: A quantitative analysis based on CiteSpace Knowledge Graph software [J]. *Ethnic higher education research*, 2019, 7 (5) : 16-24. DOI: 10.14045 / j.carol carroll nki rhen. 2019.05.004.
7. Ming, Guo Ruiying. Visualization analysis of domestic home-school cooperation research based on CiteSpace [J]. *Teaching Research*, 2017, 40(06): 27-33.
8. Zhang Yongan, Geng Zhe, Wang Yanni. Systematic classification of regional science and technology innovation policies in China: a study based on Zhongguancun data [J]. *Journal of Systems Science*, 2016, 24(02): 92-95. (in Chinese)
9. Xie Mei, Chen Wenjun. [17] The Knowledge graph of MOOCs research in China: Hot spots, current situation and Trend analysis based on CiteSpace [J]. *Journal of Southwest University for Nationalities (Humanities and Social Sciences Edition)*, 2014, 42(01): 229-235.
10. Ma Haojun. Based on the analysis of citespace new energy automotive research knowledge map [J]. *Journal of modern business*, 2019 (3) : 40 and 42. DOI: 10.14097 / j.carol carroll nki. 5392/2019.03.021.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

