Proceedings of the 4th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2025.30098

Research Hotspots and Frontiers of China's Public Data Resources—A Visual Analysis Based on CiteSpace Knowledge Graphs

Wenjing Dai

School of Accounting, Shandong Technology and Business University, Yantai, China
1758467694@qq.com

Abstract. As one of the factors of production in the new era, data resources have exerted a significant impact on society. In 2022, the Chinese government issued the "Twenty Measures on Data" policy, which clearly demonstrates the country's determination to develop data resources. As an integral part of data resources, the development of public data is an indispensable link in releasing data productivity, and plays an irreplaceable role in promoting social development and improving public efficiency. This study selects relevant literatures from the China National Knowledge Infrastructure (CNKI) database and uses the CiteSpace bibliometric tool to explore the current development status, key concerns, and future research priorities of public data in China. The research finds that public data has become a hot topic in the era of big data. Keyword clustering analysis shows that concepts centered on public data, such as data governance and digital economy, have become hot issues, while also triggering scholars' reflections on the authorized operation of public data and social responsibility.

Keywords: Public Data Resources, Visual Analysis, CiteSpace

1. Introduction

Data has become a productive force in the new era, and the exploration of its value has attracted extensive attention. As a key type of data, public data contains great potential and can bring significant convenience to the country, society, and the public. In 2022, the State Council of China issued the Opinions on Establishing a Basic Data System to Better Play the Role of Data Elements, which clearly states that public data, with its characteristics of high authority, universality, and fundamentality, plays a prominent enabling role in fields such as government governance, economic development, and common prosperity, and is an important support for advancing the basic data system. General Secretary Xi Jinping has also publicly emphasized the importance of public data utilization on multiple occasions, promoting extensive attention to this field among domestic scholars. Based on literatures from CNKI, this study uses CiteSpace software to construct knowledge graphs and conduct visual analysis, aiming to sort out the research hotspots and evolution trends in the field of China's public data resources and provide references for subsequent research.

DOI: 10.54254/2754-1169/2025.30098

2. Research methods and data sources

2.1. Research methods

This paper uses CiteSpace 6.3.R1 software to conduct bibliometric analysis on information such as the number of published papers, authors, and keywords of literatures related to China's public data resources. Based on the index maps generated by the software, it sorts out and summarizes the status of relevant research.

2.2. Data sources and research process

The data of this study are collected from the CNKI Academic Journal Database, with "public data resources" as the keyword and academic journals as the literature type. To ensure the authority and representativeness of the literatures, this study selects literatures from Peking University Core Journals and CSSCI journals. The publication time of the literatures is limited to the period from 2019 to August 2025. A total of 331 literatures are selected after screening.

Based on the above 331 literatures, the node types in CiteSpace are set to "Authors" and "Keywords" for analysis; through co-occurrence, clustering, and burst analysis, keyword knowledge graphs are generated.

3. Bibliometric analysis

3.1. Analysis of the number of published papers

A trend chart of the number of published literatures related to China's public data resources research is drawn based on the 331 screened literatures. The number of published papers in different stages reflects the degree of attention paid by the academic community to public data resources.

As can be seen from Figure 1, the number of published papers on China's public data resources can be roughly divided into two phases:

Phase 1: 2019–2020. The period from 2019 to 2020 was the embryonic stage of this research. Although the concept of "public data" was first proposed in 2015, it was often confused with concepts such as "government data" at that time. Relevant research did not form a certain scale in authoritative journals until 2020.

Phase 2: 2021–August 2025. In 2021, the Outline of the 14th Five-Year Plan for National Economic and Social Development of the People's Republic of China and the Long-Range Objectives Through the Year 2035 pointed out that public data services should be promoted to be fully integrated into the public service system, and pilot models of authorized operation of government data should be actively explored [1,2]. Since then, the usage of "public data" and "government data" has been standardized, and relevant research has flourished accordingly. This development directly responds to the call of the 2022 Opinions on Establishing a Basic Data System to Better Play the Role of Data Elements and national top-level strategies, making the development and utilization of public data a key measure to implement the "Twenty Measures on Data" and build a Digital China. Therefore, relevant research reached a historical peak in 2023–2024. More than half of 2025 has passed, and the number of published relevant literatures has exceeded that of the same period last year; it is expected that the research on public data resources will reach a new peak by the end of 2025.

DOI: 10.54254/2754-1169/2025.30098

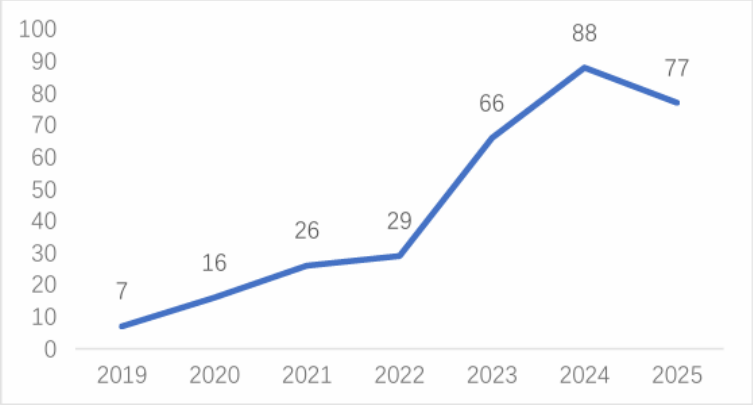

Figure 1. Trend of the number of published papers (2019–August 2025)

3.2. Author analysis

According to Table 1, a research group has initially formed around "public data resources". Ten authors, including Ma Haiqun and Zhang Huiping, have published a relatively large number of papers, with research themes focusing on data sharing and opening-up. At present, there are two major cooperation networks: one is the research on "public data security" represented by Ma Haiqun's team, and the other is the research on "public data governance" led by Hu Yefei. In general, the current research on "public data resources" has initially formed a cooperation network, but its structure is relatively loose, the connection between high-frequency authors is weak, and the research directions are relatively concentrated. Most researchers have not yet formed a systematic research system and core academic leadership. This phenomenon stems from disciplinary barriers; it is necessary to strengthen the integration of disciplines such as computer science and economics & management, and build a closer collaborative network through interdisciplinary courses and joint research.

Figure 2. Knowledge graph of high-frequency authors in public data resources research

DOI: 10.54254/2754-1169/2025.30098

Table 1. Information of authors in public data resources research

Author	Frequency	Research Theme
Ma Haiqun	7	Public Data, Data and Information Resources
Wang Jin	6	Digital Transformation, Digital Economy
Zou Chunlong	6	Public Data, Data Credit
Zhang Huiping	5	E-Government, Data Sharing
Shen Bin	3	Data Element Market, Digital Government
Zheng Lei	3	Data Management, Informatization Standards

4. Distribution of research hotspots on China's public data resources

4.1. Keyword co-occurrence analysis

Keywords centrally reflect the core content of literatures; their occurrence frequency is represented by the size of nodes, the connections between nodes reflect the co-occurrence relationship of keywords, and the thickness and structure of the connections jointly reveal the strength of the correlation [3]. The high-frequency keywords are shown in Table 2.

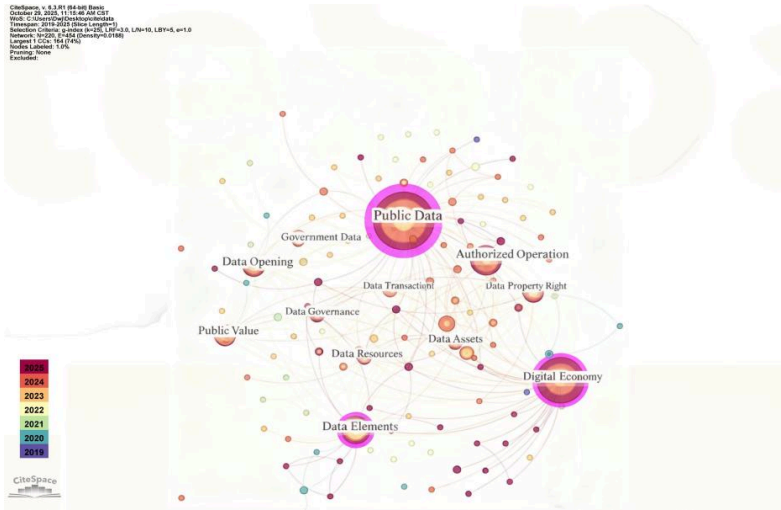

Figure 3. Keyword co-occurrence knowledge graph

DOI: 10.54254/2754-1169/2025.30098

Table 2. Information of high-frequency keywords

Frequency	Centrality	First Appearance Year	Keyword
114	0.6	2019	Public Data
67	0.27	2022	Data Elements
30	0.04	2022	Authorized Operation
28	0.24	2020	Digital Economy
22	0.06	2019	Data Opening-up
19	0.08	2020	Data Governance
15	0.04	2023	Data Property Rights
11	0.05	2020	Government Data
11	0.07	2021	Data Resources

The keywords can be divided into the following categories: first, the category of basic data and core elements, such as Public Data and Data Assets; second, the category of data governance and operation, such as data authorization, operation, and governance; third, the category of data application and value embodiment, such as Digital Economy.

In Table 2, centrality refers to the degree to which each keyword acts as an "intermediary", representing the number of times the keyword serves as a hub for multiple other keywords. The higher the centrality, the greater the mediating role of the keyword and the more representative it is in research. Several keywords with relatively high centrality in the table are "Public Data", "Data Elements", "Authorized Operation", and "Digital Economy".

4.2. Keyword clustering analysis

In CiteSpace, when the modularity $Q > 0.5$ and the silhouette value $S > 0.7$, the clustering structure is generally considered significant and the clustering reliability is high. The clustering graph obtained in this study has $Q = 0.5148$ and $S = 0.8395$, which meets the effective clustering standards, indicating that the keyword clustering results have practical analysis value [3]. Keyword clustering analysis formed a total of 6 clusters, including #0 Authorized Operation, #1 Digital Economy, #2 Data Elements, #3 Data Governance, #4 Public Value, and #5 Big Data. The research themes can be summarized into three categories: first, the mechanism of action of data elements in the digital economy; second, the exploration of data authorized operation models; third, data governance and standardized utilization based on Big Data. This reflects that the research on public data resources has been highly popular in recent years and shows an obvious practical orientation.

DOI: 10.54254/2754-1169/2025.30098

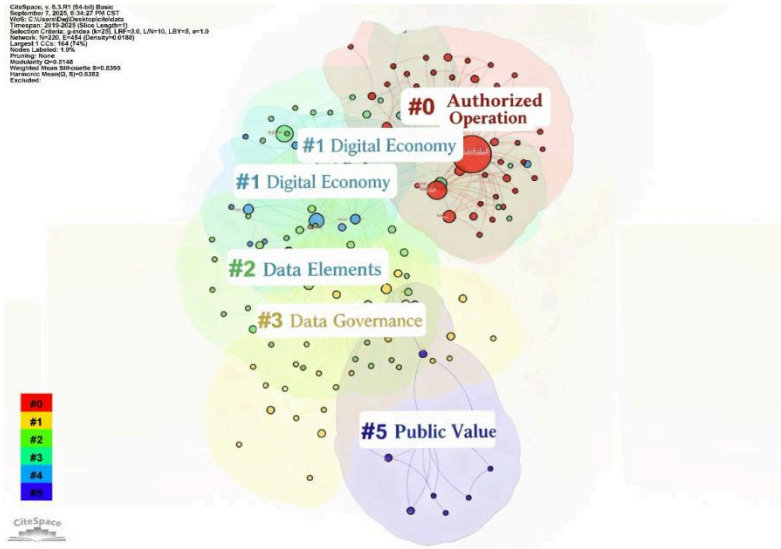

Figure 4. Keyword clustering knowledge graph

4.3. Keyword time series graph analysis

Time series analysis clearly presents the evolution path of each research hotspot. As shown in Figure 5, since the rise of "Data Elements" in 2019, key directions such as "Data Assets" and "Data Transaction" have been gradually derived; the basic concept of "Public Data" has continued to expand, giving rise to new topics such as "Authorized Operation". At the same time, "Social Responsibility" has been deeply integrated with data governance and enterprise innovation, promoting the formation of a research context for standardized data application. The research focus of the early hotspot "Big Data" has gradually shifted to new fields such as "Data Governance" and "Data Registration", while clusters such as "Digital Economy" and "Urban Governance" have been gradually formed along with the marketization process of data elements, which is in line with the overall development logic.

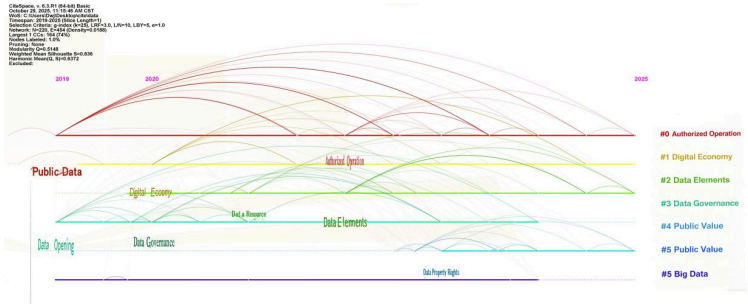

Figure 5. Keyword time series knowledge graph

4.4. Keyword burst analysis

To capture research frontiers, this paper analyzes burst words (keywords with a significant increase in frequency), with the parameter γ set to 0.5 [4]. The results show that a total of 8 burst words were generated between 2019 and 2023 (Figure 6), which can be divided into two phases according to their time distribution.

DOI: 10.54254/2754-1169/2025.30098

Phase 1 (2019–2020) saw the emergence of burst words such as "Big Data" and "E-Government". At this time, the research was still in the theoretical exploration stage, and in practice, it mainly focused on using Big Data technology to improve the efficiency of fields such as public resource transactions [5]. Scholars such as Cai Shu believe that the application of Big Data in the public sector requires the government to transform into a data-driven one and systematically reshape decision-making and business processes, as it has become the key to national core competitiveness and local breakthrough development [6].

Phase 2 is 2021–2024, with burst words including "Library", "Public Culture", and "Data Finance". Based on the research foundation of previous years, the research direction in this period has gradually deepened. From the perspective of research content, the research focus not only pays attention to data opening-up itself, but also begins to multi-dimensionally examine the path of public data empowering enterprises and the internal connection of its social value realization (Zheng Zhiqiang, 2024). The promoting effect of public data opening-up on the entry of new enterprises is mainly reflected in non-state-owned enterprises, the service industry, and cities with a high Internet penetration rate [7].

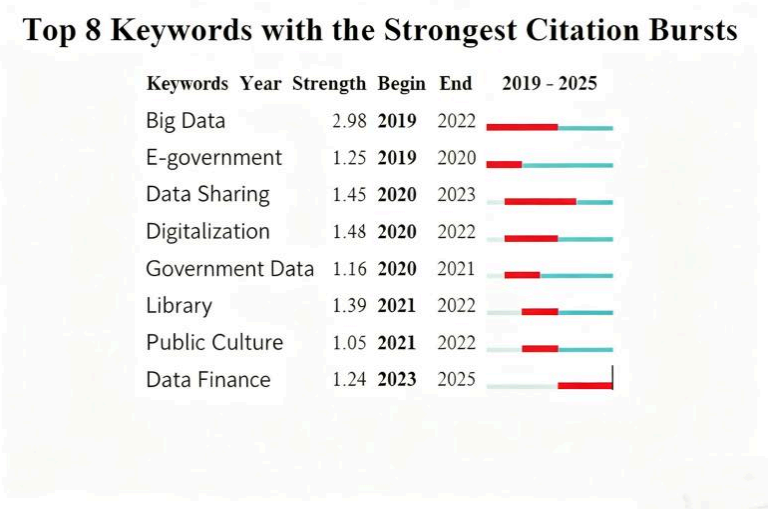

Figure 6. Keyword burst knowledge graph

5. Summary of research hotspots on China's public data resources

5.1. Analysis of the driving factors for the formation of research hotspot characteristics

The evolution of research hotspots in public data resources is the result of the joint action of internal and external factors, which is specifically reflected in the following three aspects:

The core driving force is following policy guidance. The characteristics of public data resources make their development highly dependent on national institutional design. From the "authorized operation" in the "14th Five-Year Plan" to the basic institutional design in the "Twenty Measures on Data", the top-level design has continuously driven the in-depth development of research hotspots from "data opening-up" to value realization fields such as "authorized operation" and "data elements".

Promoting value release is the internal driving force. Research hotspots accurately correspond to the paths for exploring the multi-level value of public data resources. First is its inherent economic value, forming a research chain of "Data Elements → Data Assets"; second is returning public data

DOI: 10.54254/2754-1169/2025.30098

resources to public value, which is reflected in the attention to fields such as "Public Culture" and the exploration of its mission in promoting social equity.

Facing governance challenges is the problem pressure. The scale and complexity of public data resources have brought practical problems such as security, privacy, standards, and fair access. The practical transformation from "opening-up" to "operation" has directly given birth to and continuously strengthened the research popularity of "data governance".

5.2. Research on "the relationship between public data resource opening-up and enterprises"

The relationship between public data resource opening-up and enterprises is a research field that has attracted much attention in the current information society. Many scholars have conducted research on this topic from different perspectives. Studies have shown that the opening-up of public data resources can drive the technological innovation and business model innovation of enterprises, not only improving their core competitiveness through technology and product research and development, but also promoting their transformation and development by exploring new business opportunities.

In addition, some studies have discussed the models and mechanisms of cooperation between public data resource opening-up and enterprises, explored the cooperative relationship between the government, enterprises, and social organizations, and how to establish effective data sharing and utilization mechanisms. However, some literatures also point out the problems faced by the opening-up of public data resources, such as data privacy protection and security risks, which have a certain impact on the data utilization and development of enterprises.

5.3. Research on "the role of public data resources in the digital economy"

The research on the role of public data resources in the digital economy is a research field of great practical significance under the background of the current digital era.

Some studies focus on the driving role of public data resources in the innovation of the digital industry, pointing out that as a key factor of production, public data resources can provide rich data support for digital technology research and development and digital product design, help digital industries such as artificial intelligence, Big Data, and cloud computing break through technical bottlenecks, give birth to more innovative digital products and services, and promote the continuous improvement of the digital industry system.

Relevant studies point out that the effective circulation and sharing of public data resources help reduce the information acquisition cost of market entities, alleviate information asymmetry, thereby promoting fair market competition, creating a more standardized and orderly environment for the development of the digital economy, and attracting more enterprises to participate in it. Other studies show that the combination of public data and traditional industry data can effectively promote the process of industrial digitalization, breed new business formats and business models such as smart agriculture and smart healthcare, open up a broader application space for the development of the digital economy, and further enhance its promoting effect on the overall economy.

5.4. Research on "the assetization and value evaluation of public data resources"

As "Data Assets" and "Data Elements" have become high-frequency keywords, the research focus has shifted from confirming the asset attribute of public data to solving the practical problems of its asset recognition and measurement. As scholars such as Zhao Zhigang have pointed out, the core

DOI: 10.54254/2754-1169/2025.30098

difficulty lies in how to bridge the gap between theory and practice and reliably measure and report it in financial statements [8]. This challenge has given birth to the diversified exploration of evaluation methodologies. At present, the current research has developed a variety of evaluation models and methods, mainly including the three-dimensional static evaluation - classified dynamic evaluation (covering three dimensions of quality, utility, and ecology) and dynamic paths designed for personal, enterprise, and public data. In addition, combined with machine learning, multiple characteristic indicators such as data capacity, market attention, data application degree, and data freshness are selected to reduce subjectivity [9].

5.5. Research analysis on "the financial performance and impact of public data resource operation"

The economic value of public data resource operation needs to be tested through its substantial impact on micro-subjects and the macro-economy.

At the enterprise level, studies by Hu Jinyan and others have found that the key to the promoting effect of public data opening-up on enterprise performance lies in its realization of operational empowerment and management empowerment [10]. This effect is particularly prominent in non-state-owned enterprises, small enterprises, and enterprises with a high degree of digitalization, indicating that public data helps to make up for competitive disadvantages and stimulate innovative vitality.

Studies at the government level have confirmed that the development level of the digital economy is the key path for public data to promote the improvement of government performance, while the activity of digital innovation and public participation also play a significant positive regulatory role [11]. Studies further found that public data policies are more effective in underdeveloped regions, highlighting their positive role in breaking development barriers and making up for resource shortcomings. Against this background, the concept of "Data Finance" has emerged, and the academic community has begun to discuss the feasibility of using the income from the authorized operation of public data as a new local financial resource and pay attention to its potential impact on the local fiscal structure and sustainability. Fu Qiang and others found that public data indirectly promotes the sustainable development of local finance through two intermediary paths: increasing income and optimizing expenditure, and this effect is more prominent in cities with a high dependence on land finance and greatly affected by tax and fee reductions [12].

However, the current research on financial performance is still mostly limited to the "black box" model, and there is insufficient research on the specific transmission path of value creation. At the same time, the unclear distribution mechanism of data operation income and the lack of an accounting and reporting system for "data assets" at the government level make the macro picture of data finance still relatively vague. It is urgent to construct a complete "operation-behavior-performance" analysis framework from the perspective of financial management.

6. Conclusion

To demonstrate the development of China's public data resources, this paper uses CiteSpace visual measurement software for analysis. Under the premise of limiting the time range to 2019–2025, an analysis was conducted on the authors, number of published papers, keywords, and research hotspots of the research theme. The conclusions are as follows: (1) In terms of research evolution, it was in the embryonic stage before 2020, and then grew rapidly driven by policies such as the "Twenty Measures on Data" and technological progress, but the loose author cooperation network

DOI: 10.54254/2754-1169/2025.30098

has restricted the research depth; (2) In terms of hotspot evolution, the research focus has expanded from "data opening-up" to mechanism construction fields such as "authorized operation" and "data governance", and extended to value realization frontiers such as "data assets" and "data finance"; (3) In terms of frontier trends, although the research has touched on the assetization path and financial impact mechanism, there are still theoretical gaps in accounting recognition and value evaluation. In the future, a "policy-technology-finance" integration framework should be constructed, focusing on breaking through interdisciplinary issues such as assetization accounting and financial performance transmission mechanism, so as to provide theoretical support for activating the value of public data.

References

[1] CSSN. Promoting the opening-up and authorized operation of public data in a coordinated manner [EB/OL]. (2024-03-28) [2025-09-08]. https: //www.cssn.cn/skgz/bwyc/202403/t20240328_5742750.shtml
[2] People's Republic of China, National People's Congress, et al. Outline of the 14th Five-Year Plan for National Economic and Social Development of the People's Republic of China and the Long-Range Objectives Through the Year 2035 [EB/OL]. (2021-03-11) [2025-09-08]. https: //baike.baidu.com
[3] Wang, J. Y., & Song, J. Q. (2023). Review of research status and hotspots of laboratory safety management in China: Visual analysis based on CiteSpace. China-Arab States Science and Technology Forum (Chinese & English), (12), 112–117.
[4] Yu, Z., & Wang, Y. (2023). Visual analysis of primary school education research in China from the perspective of knowledge graph. Journal of Mudanjiang College of Education, (10), 33–37, 74.
[5] Huang, S., Chen, H., & Lu, K. C. (2016). Research on the advantages of realizing information sharing in public resource transactions using big data. Enterprise Science and Technology & Development, (2), 3.
[6] Cai, S., & Pang, C. Q. (2016). Innovation of government governance thinking in the era of big data. Hangzhou Science and Technology, (04), 23–26.
[7] Zheng, Z. Q., & He, J. L. (2024). Awakening "sleeping" data: Public data opening-up and new enterprise entry. Foreign Economics & Management, 1–16. Retrieved June 30, 2024.
[8] Zhao, Z. G. (2025). Core difficulties and promotion strategies of public data capitalization. Chinese Public Administration, 41(05), 149–152. 2025-05-014.
[9] Li, Y., & Yao, L. H. (2025). Optimization of public data value evaluation model: Empirical analysis based on machine learning algorithm. Journal of Taiyuan University of Technology (Social Sciences Edition), 43(03), 120–127, 136.
[10] Hu, J. Y., Yu, L., & Liu, Y. W. (2024). Does digital government construction improve enterprise performance? From the perspective of data element opening-up. Dongyue Tribune, 45(06), 109–122. 2024.06.013.
[11] Wu, X. L., Zheng, J. E., & Ma, S. Z. (2025). Public data opening-up and digital government performance: Analysis based on multi-period difference-in-differences model. Research on Economic Development, (03), 18–33. 0.2025-03-002.
[12] Fu, Q., & Hao, Y. L. (2025). Research on the impact of public data opening-up on the sustainability of local finance. Finance and Trade Research, 36(07), 75–93. https: //doi.org/10.19337/j.cnki.34-1093/f.2025.07.006