IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 25 March 2024, accepted 27 May 2024, date of publication 31 May 2024, date of current version 22 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3408093

RESEARCH ARTICLE

Content Summary and Research Visual Analysis of Data Product in China Based on CiteSpace

YANNI LIANG [1,2,3], JIANXIN YOU [1], RAN WANG [4], AND SHUO HAN [2]

[1] School of Economics and Management, Tongji University, Shanghai 200092, China
[2] School of Economics and Management, Beibu Gulf University, Qinzhou 535011, China
[3] Beibu Gulf Marine Development Research Center, Qinzhou 535011, China
[4] SkyCloud Technology Company Ltd., Guangzhou 510308, China

Corresponding author: Jianxin You (yjx2256@vip.sina.com)

This work was supported by Guangxi Philosophy and Social Sciences Research Project under Grant 23CYJ022.

ABSTRACT Data products play an important role in driving the digital transformation of enterprises, establishing the market of data elements, and fostering the sustainable development of the economy. Based on the research of data products in Chinese academic circles, this paper summarizes and defines the concept of data products, systematically collects Chinese policies related to data elements, and forms a literature review of data products. With the CiteSpace bibliometric analysis software, it conducts an analysis of authoritative literature on data products both domestically and internationally up to 2023 by using the relevant research literature from China National Knowledge Infrastructure and Web of Science databases as data sources. It delves into the author institution, topic evolution and research trend within the realm of data product research. The results demonstrate that data products are multi-dimensional and complex, with their evolution significantly influenced by policy and legal frameworks. Since 2010, both domestic and international research on data products has seen rapid growth. The Chinese academia has gradually expanded from focusing on data market expansion and service system to meteorology and earth sciences, while the international research has involved the application of data products in soil moisture, UV scattering, precipitation monitoring, remote sensing technology, geoclimate and other fields. The research recommends strengthening interdisciplinary, cross-field, cross-team and industry-academia-research-application cooperation and international exchanges. With the development of new technologies such as artificial intelligence, it should pay more attention to the quality and performance evaluation of data products, data security and privacy protection, and the value co-creation and income distribution of data products. This paper is aimed to promote the application of data products in social management and business fields, and maximizing the value and potential of data products.

INDEX TERMS Data product, visual analysis, CiteSpace, definition, policy.

I. INTRODUCTION

With the continuous advancement of digital technology and the deepening of digital transformation, data and its derived value play an increasingly important role in promoting global economic development. According to the *Digital China Development Report 2022* [1] released by the Cyberspace Administration of China, Chinese data output reached 8.1ZB in 2022, up 22.7 percent year on year, accounted for 10.5 percent of the world's total data output and ranked second in the world, and 48 data trading institutions have been established nationwide. Data, data element, data resource, data asset and data product constitute different levels of data management and utilization in the value chain of data products. The transformation process from data to data products reflects the core operation mechanism of the digital economy, which is to transform data into economic value through technology and innovation. Participation in data trading primarily involves five types of entities: data supplier, data merchant, data exchange platform, data demander,

The associate editor coordinating the review of this manuscript and approving it for publication was M. Shamim Kaiser.

TABLE 1. Definitions of data products in chinese academia.

Author	Definition
Song and Qin [4]	There are four roles in the data product manufacturing environment: data provider, data producer, data consumer, and data manager.
Yu and Zhang [5]	Refers to data products and services that have been processed through crawling, reformatting, cleaning, encryption, etc, such as data sets and information services derived from data sets.
Pei [6]	A data set as a product, or an information service derived from a data set.
Xiong and Tang [7]	It is the manifestation and organizational result of information integration. It is stored on data carriers such as computers and cloud disks, and can be connected, integrated and associated with the behavior trajectory and associated information of a specific object, which has strong analytical value.
Ye et al. [8]	Single-type data products such as digital songs, videos, images, e-books, and online novels have basic characteristics including unified format and capacity, complete content, and replicability. Their supporting systems usually include data platforms, user interfaces, pricing rules, copyright protection management mechanisms, and data infrastructure.
Li et al. [9]	Through business digitization, enterprises use data platforms to collect and transmit data, store and integrate it into valuable digital assets. They accurately classify this data and apply it to various users and fields. Data applications include data calls, sharing, and decision-making, forming diverse data products.
Ji [10]	Data products refer to the deep processing and processing of data sets to form an intelligent decision, which can be used as the basis for product upgrading or enterprise marketing plans.
Ouyang and Gong [11]	Data products are standard products or services after the commercialization of data resources, the carrier of data value and asset, and the main transaction object of data factor market, containing rich application value and exchange value.
Ouyang and Du [12]	In the data factor market, data products mainly include data package, data API, data report and data service according to the degree of processing and transmission technology.
Li [13]	Network operators obtain original data via legal means, apply certain algorithms to this data, and generate derivative data with exchange value and technical feasibility after in-depth analysis, filtering, refining, integration, and desensitization. The data product has certain expression, which can offer users predictive, indexing and statistical services via the content of the derivative data.
Gao and Ran [14]	Tradable and circulating data products include: (1) Data products of knowledge production factors, referring to reusable raw data after processing; (2) Data products of knowledge production tools, encompassing data analysis methods or technical tools rather than data elements themselves; (3) Knowledge service data products, as an element of knowledge production. Once a machine can learn rules from large historical data sets and form models, it can generate reasonable predictions or insights based on input data state.
Gao [15]	Data is a general upper concept, and data sets with commodity attributes are formed by data processing and processing activities as data products.
Huang et al. [16]	Data products can be divided into four categories: original data sets, desensitized data sets, modeled data and artificial intelligence data.
Wang [17]	It refers to the product formed by processing the data obtained by the enterprise.
Gao [18]	Data is divided into data resources (or "data elements") in original form and output data products used in data processing. Data products are essentially the ultimate realization of data value.
Gao [19]	Data products are derivative data generated based on the development rights of network operators, and their development is divided into small data and big data stages. (1) Data products in the small data stage are not directly related to digital labor, and in this case are based on small data and reflected as knowledge products and consulting services. (2) Data products in the big data stage are products of digital labor, and in the big data stage develop new use values and form new product forms based on large-scale data processed by workers using intelligent labor tools in production.
Liu [20]	The object of data property rights is not data itself, not merely a collection of big data, nor a single piece of information, but rather a product of specific information content processed via big data analysis technologies. This product can be used to solve specific problems and meet specific business needs, which can be called the data product.
Liu [21]	Data products can be divided into broad and narrow senses. (1) In the broad sense, a data product refers to any commodity or service built on binary code and existing within a computer information system, such as online game services, electronic currency, and virtual property. (2) In the narrow sense, a data product refers to derivative data with exchange value and technical feasibility obtained through in-depth analysis, filtering, refining, integrating, and desensitizing original data via algorithm rules. It provides users with prediction, indexing, statistical, and other services.
Li and Gao [22]	Data product refers to a new data form that enterprises desensitize, process and mine data sets, which can provide prediction for enterprise decision-making.
Huang [23]	Data products are processed by people, while data resources exist without being processed by users. (1) for data producers, all information collection objects are their data resources, which become data as products after being converted into usable data. (2) For data users, all data resources and data products can be regarded as data resources, no matter they are used to produce other products, labor services or direct consumption.
Sun [24]	Data products refer to derivative data and data derivative products formed by substantial processing of original data, belonging to data resources. Original data property rights motivate "data collection labor", while data product property rights motivate "data processing labor".

Note: The above table counts the relevant research literature retrieved by Chinese Social Sciences Citation Index, Science Citation Index Expanded (SCIE) and Social Sciences Citation Index (SSCI).

and assessment service organization [2]. Data products and service providers mainly include Internet enterprises with massive self-owned data, IT enterprises with advanced big data platform construction technology and data analysis and processing technology, third-party service companies relying on data analysis and processing tools to provide data services according to users' needs, and big data exchanges specializing in data matching transactions [3]. The current definition of data products has not been clear. DJ Patil, the chief data scientist in the United States, defines data product as products that achieve a business goal through the use of data. Zhejiang University's *Data Product Transaction Standardization White Paper 2022* defines data product transactions as the act of the data demander obtains data products by the data supplier by the money purchases or exchanges within the scope of national laws and the guarantee of a safe transaction environment and transaction compliance supervision. Additionally, the concept and definition of data product have also been explored in Chinese academia.

The definition of data products in Chinese academia reflects the multi-dimensional and complex characteristics of its concept, involving data quality, pricing, legal protection, market mechanism and information technology. (1) The role and function of data products: most scholars believe that data products are not only a collection of data, but also product with practical application value formed through specific processing and processing. For example, data products can be data sets that have been cleaned, encrypted and deeply analyzed, or information services provided by these data sets. (2) Market and value of data products: Some scholars have emphasized the role and value of data products in the market. Data products are regarded as standard products or services after the commercialization of data resources. They not only carry the value and asset of data, but also are the main transaction objects in the data market. (3) The legal status and protection of data products: some scholars pay attention to the legal status and protection of data products. It involves the legal acquisition, processing and use of data collections, as well as the legal issues of data products in terms of intellectual property rights and personal information protection. (4) Relationship between data products and technological development: The development of technology, especially in the field of big data and AI, has a significant impact on the definition and value of data products. Data products can be data sets that have been processed through advanced algorithms and technologies, which can enhance the analytical and predictive capabilities of data products. Based on the previous definitions, this paper summarizes that data products can be defined as processed and processed data resources, which transform raw data into information products with commercial value and practical application potential, including data sets, analysis reports, prediction models, etc. They not only reflect the value and asset of data, but also the main transaction object of data market, and are legally involved in legal acquisition, use and protection. The application scope and functional definition of data products will continue to expand and evolve with the development of technology. After discussing the definition of data products, this paper sorts out the closely related data element policies. As the basis of data products, their management and specification directly affect the quality, security and business value of data products. Therefore, sorting out and analyzing the policies related to data elements are not only conducive to understand the legal and business environment of data products, but also crucial to ensure the compliance, innovation and market competitiveness of data products. The following table shows the national policies related to data products and data elements released in China in recent years.

The above policies collectively show China's attention and emphasis on data products and their management, market development, standardization, asset management and legal protection, which are mainly reflected in five aspects. (1) Unified and standardized data management regulation, which emphasizes the establishment and improvement of data management system, involving data quality improvement, ownership definition, open sharing and transaction

TABLE 2. Policies related to data products issued in china in recent years.

Release Date	Policy Name	Related Content
March 2020	*Opinions on Building a More Complete System and Mechanism for Market-Based Allocation of Factors* [25]	Explore the establishment of a unified and standardized data management system to improve the quality and standardization of data and enrich data products.
May 2020	*Opinions on Accelerating the Improvement of the Socialist Market Economic System in the New era* [26]	Accelerate the development of the data element market, establish the list management mechanism for data resources, improve standards and measures for defining data ownership, opening up and sharing data, and trading and circulating data, and utilize the value of social data resources.
September 2020	*Opinions on Leading the Accelerated Development of New Consumption with New Forms and Models* [27]	Enhance the commercial level of consumption information and data sharing, and better provide computing power resource support and preferential services for enterprises.
December 2020	*Instruction Opinions on Accelerating the Construction of a National Integrated Big Data Center Collaborative Innovation System* [28]	Improve the new big data integration trading mechanism covering different data development levels, including raw data, desensitized data, modeled data and AI data.
January 2021	*Action Plan for Establishing a High-standard Market System* [29]	Establish basic systems and standard specifications for property rights, transaction circulation, cross-border transmission and security of data resources, and promote the development and utilization of data resources.
March 2021	*The 14th Five-Year Plan and 2035 Long-range Goal Outline for National Economic and Social Development of the People's Republic of China* [30]	Establish and improve data property rights trading and industry self-discipline mechanisms, and foster standardized data trading platforms and market players.
November 2021	*"The 14th Five-Year" Big Data Industry Development Plan* [31]	Promote the cultivation of a standard compliance evaluation system that covers data product evaluation, data resource planning, data governance implementation, data asset evaluation, data service capabilities, etc.
December 2021	*Notice on Issuing the Overall Plan for Pilot Comprehensive Reform of Factor Marketization Allocation* [32]	Explore data-centered product and service innovation in key areas such as finance, health care, electricity and logistics, support the establishment of unified technical standards and an open innovation ecosystem, and promote the commercial data flow, cross-regional data interconnection, and the integrated application of government-enterprise data.
December 2021	*Notice on Issuing "the 14th Five-Year" Digital Economy Development Plan* [33]	Encourage market forces to tap the value of commercial data, promote the productization and servitization of data value, vigorously develop professional and personalized data

TABLE 2. ***(Continued.)*** **Policies related to data products issued in china in recent years.**

Release Date	Policy Name	Related Content
		services, promote the deep integration of data, technology and scenarios, and meet the data demands of various fields.
March 2022	*Opinions on Speeding up the Establishing of a National Unified Large Market* [34]	Accelerate the cultivation of data element market, establish and improve basic systems and standards for data security, rights protection, cross-border transmission and management, transaction and circulation, open sharing and security authentication, conduct in-depth surveys on data resources, and promote the development and utilization of data resources.
December 2022	*Opinions on Establishing Data Basic System to Better Play the Role of Data Elements* [35]	Support third-party institutions and intermediary service organizations to strengthen the formulation of data collection and quality assessment standards, promote the standardization of data products, and develop industries such as data analysis and data services. Establish a revenue distribution system for data factors that reflects efficiency and promotes fairness.
February 2023	*Overall Layout Planning for the Construction of Digital China* [36]	Unleash the value potential of commercial data, accelerate the establishment of a data property rights system, carry out research on the valuation of data assets, and establish a participate and distribution mechanism for data elements according to the value contribution.
December 2023	*Notice on Issuing the "Data Elements ×" Three-year Action Plan (2024-2026)* [37]	By the end of 2026, a relatively comprehensive data industry ecosystem will have been established, leading to a notable enhancement in the quality and efficiency of data products and services. The data industry is expected to experience an annual growth rate exceeding 20%. Twelve key industries and sectors, such as industrial manufacturing, modern agriculture, trade circulation, transportation, financial services, scientific and technological innovation, cultural tourism, medical and health care, emergency management, meteorological services, urban governance, and green and low-carbon initiatives, will be prioritized for driving the multiplication effect of data elements to release the data element value.
December 2023	*Notice on Issuing the Implementation Plan for Digital Economy to Promote Common Prosperity* [38]	Strengthen the development and utilization of public data resources, promote efficient and compliant use of data, cultivate data element enterprises, prosper the data element market, and further activate the dividend of data elements.
December 2023	*Notice on issuing the Instruction Opinions of Strengthening the Management of Data Assets* [39]	Implement the rights separation requirements of data resource ownership, data processing and usage rights, and data product operating rights, and accelerate the construction of a scientifically classified data asset property rights system.

circulation. (2) Development of data factor market, which focuses on the cultivation and development of data factor market, including the establishment and improvement of data security, rights protection, transaction circulation and other systems and standards. (3) Data property rights and transaction mechanisms, which promotes the establishment of data resource property rights and transaction circulation mechanisms, including cross-border transmission and data security management. (4) Standardization of data products and services, which encourages the standardization and quality improvement of data products and services, involving data product evaluation, data governance implementation, etc. (5) Data asset management and value realization, which focuses on the management and value realization of data assets, including data asset valuation research and the establishment of data property rights system.

II. RELEVANT REFERENCE ANALYSIS

In this paper, data product is used as a search term to analyze the research content and progress of data product in Chinese academia through the authoritative academic platform China National Knowledge Infrastructure (CNKI) and Web of Science (WOS), which involves the database of Peking University Core, SCIE and SSCI. It can be summarized into the following four categories.

A. FROM THE PERSPECTIVE OF CONCEPTUAL DEFINITION AND RIGHT-INTEREST DETERMINATION OF DATA ELEMENTS

Duan [40] proposed that data products possess characteristics such as low replication cost, significant value added through data integration and aggregation, and non-exclusivity allowing for simultaneous use by multiple parties. Li [13] defined data products, and insisted that data products should belong to property rights, with exclusive and clear connotation of rights, and have the right to possess, use, profit and disposal. Gao [15] proposed a classification of data into original data and derivative data according to its dynamic value cycle, and posited that data products should be divided into pooled data and derivative data according to the difference of labor paid in producing data products. The former is an intermediate product that does not generate new commercial and social value. Ji [10] posited that data property rights are defined by the exclusive rights of data controllers to control, use, transmit, and dispose of data products, and proposed that data value generation mechanism consists of four key links: collection, storage, processing and utilization. Wang [17] believed that although data products are the property of enterprises, if the R&D and design of data products involve personal information, the legal processing of personal information should be the basic premise. Pang [41] proposed that the acquisition method of data utilization rights should be clarified, the property rights and interests of data sets should be allocated according to the contribution degree of data-related stakeholders with the help of the right confirmation function of private law, and data processing should be used to form data products with both property and personality attributes. Huang [23] posited that the property rights relationship between data resources and data products is complex and diversified, and proposed that in ''big data'' products, the contribution of information source owners is shared by consumers and merchants. He suggested that the state could collect information resource tax for the provision of national or regional public goods. Based on the demonstration of digital labor, Gao [19] believed that the labor process of transforming data into data products with use value mainly includes three steps: data resource utilization, data production factor utilization and data development. The first step is to obtain behavioral data through the selection of data collectors and the setting of collection equipment. Then, the data analyst cleans and screens the data to ensure its integrity and effectiveness and uploads it to the processing system. Finally, the data engineer uses algorithms to process the data to complete modeling, testing, iteration and optimization. Li and Gao [22] suggested to construct a limited exclusive right to balance the relationship between the flow and the property rights protection of data products. Liu [20] proposed that the producers of data products can hold, operate and manage data products and obtain profits. In order to limit the degree of property rights of data products, the producers' rights should include the rights of acquisition, use and public dissemination, and the data sources should enjoy the rights of copying, accessing and correction of corresponding data to ensure fair returns. Liu [21] believed that data products are property interests based on computing power and have high economic value, which can be valued according to the relative value of commercialized computing power. And the subject of rights belongs to the manufacturer of the product. Gao [18] proposed that data products are essentially the ultimate realization of data value, which needs to be allocated to data holders with stable and certain exclusive property rights. Xiao [42] proposed that the government data authorization operation belongs to one of the PPP models of government-private capital cooperation, and its special legislation needs to carry out specific system design around the authorization operation link, data products and service market-oriented provision link respectively. Sun [24] defined and distinguished data rights objects such as data, data resources, data products and data collections as the prerequisite for data right confirmation, and proposed that data commodities and data assets are different from data products and are not data. data commodities themselves do not create value but only releases value, and the evaluation of data assets do not create value but only discover the value.

B. FROM THE PERSPECTIVE OF CONTENT IDENTIFICATION, DEVELOPMENT DESIGN OF DATA PRODUCTS

Liao et al. [43] divided geoscience data into three categories based on the characteristics of geoscience data: geoscience data products based on attribute data, geoscience data products based on remote sensing information sources and data products based on vector maps. They proposed that geoscience data sharing and publishing platform should have basic functions such as user management, data catalog query, metadata management, data query and browsing, and data download. Ma et al. [44] introduced that NASA's construction of Earth observation science data product generation link includes three basic stages: user demand and technical level feasibility analysis, software platform and algorithm development, and product generation and service. An and Huang [45] expounded the source and access of digital data products from the concept of library data service, and analyzed the content level and knowledge and skills requirements of data service. Wu et al. [46] discussed and designed a WikiSensing platform that can effectively explore the value of sensor data. The platform includes data storage, data query, application development and product market, and supports the whole life cycle from sensor data to data products. Wu and Zhang [47] systematically sorted out the existing methods to obtain valuable data products from remote sensing observation data, and summarized and analyzed the advantages of target recognition and parameter extraction. Tian et al. [48] systematically sorted out domestic and foreign service providers that provided ESG data products and their products and services. Liu et al. [49] argued that strong consistency did not meet the data management requirements of high scalability, high performance, high fault tolerance, high scalability and high economy in the era of big data, and the development of data products should focus on the integration of data science theories to improve product upgrading and user experience. Ye et al. [8] focused on the system-level design of big data product and systematically studied the pipeline including containerization, circulation, pricing, protection and etc. by referring to information media such as books. They also proposed that data products could be traded on the data market only if their rights were registered and protected by law. Guo and Zhu [50] summarized the spatialization methods, auxiliary data and main data products of social and economic statistical data. They also outlined trends in the spatialization method improvement, new auxiliary data exploration, multi-source data comprehensive utilization, and the high spatio-temporal resolution and high precision data product research and development. Wan and Gu [51] used the multi-case secondary data analysis method to analyze the data products of the literature and information industry and their structure from the perspective of data supply chain. Gao [52] explored the construction stage and construction method of industrial big data products of Sinopec Sharing Service Company from the perspective of shared big data application.

C. FROM THE PERSPECTIVE OF VALUE CREATION AND REVENUE DISTRIBUTION OF DATA PRODUCTS

Xiong et al. [53] proposed that one of the biggest difficulties in the data product pricing lied in the multi-directional uncertainty of the value of data products, which is embodied in the different understandings and recognition of the buyers and sellers on the expected value of data utility, the degree of data cleaning and the uniqueness of data before and after data trading. Luo and Xing [54] introduced the optimistic value theory into neutrosophic variables for observing the status of player, and constructed different pricing models of satellite image data products under the Bertrand and Stackelberg game scenarios. Luo et al. [55] investigate the pricing problems of satellite image data products in a neutrosophic fuzzy environment, and proved that it is generally more favorable for customers when two providers play Bertrand game. Zhao and Sun [3] proposed that the value of big data transaction is composed of the characteristic price of big data transaction, the excess return of producers and traders, and the return of buyers. They suggested that big data products adopt the trial price mechanism based on supply and demand to ensure that producers, traders and buyers all share the value of big data. Xiong and Tang [7] proposed that due to the more obvious characteristics of data such as replicability, low marginal cost, integrability and value uncertainty, the transaction mode and seller strategy of data products are more affected by application scenarios, buyer heterogeneity and market structure. Ouyang and Du [12] believed that the total cost of data, data value and application scenarios are important factors to determine the price of data products, and the value of data elements can be calculated according to typical application scenarios, and combining with the integrity, accuracy, hierarchy, coordination and heterogeneity of data elements. Gao and Ran [14] posited that data transaction is a demand-driven market rather than a supply-determined market, and data transaction platform should build a safe and reliable data circulation channel around the realization of data product value to achieve reusable data service domain. By referring to the price generation mechanism of securities market, Huang et al. [16] proposed a price formation mechanism of data products covering four types of subjects including data seller, data buyer, data exchange and third-party institution, as well as a price generation path combining quote-valuation-negotiation. Ouyang and Gong [11] proposed that data factor pricing mechanism is the process of data factor value being externalized into data product price along the evolution path of value formation - price discovery - bidding and transaction. Li et al. [9] took litchi industry as a case study and proposed that data value chain includes business data, data platform, data product, product market and product monetization, and posited that data had become a new profit source of agricultural industry. Yu and Huang [56] analyzed the data transaction process of three different stakeholders including end users, data suppliers and data service providers, and proposed three different data pricing models of centralized pricing, decentralized pricing and revenue sharing based on closed-loop data supply chain. Guo et al. [57] constructed a duopoly competition game model, divided the game process between data trading platforms into two stages of quality decision-making and price setting. They analyzed the relationship between data trading sampling strategy and data product quality and platform revenue in equilibrium. Wang et al. [58] believed that data pricing is based on the results of valuation and equilibrium between multiple parties, and different data subjects have different valuations for data products. Pan et al. [59] proposed that the quotation-valuation-pricing mechanism of data product trading is the service capability of the platform, which is the internal requirement for realizing and expanding data trading on the exchange. Combining with the characteristics of data products, Guo et al. [60] analyzed the transaction behaviors between the data trading platform and potential data demanders, and constructed a revenue model for both parties of data trading. Wang et al. [61] proposed that the final value of data products was closely related to prediction accuracy, prediction robustness, application scenarios, market environment, data credit, data input cost, etc. And the benefit distribution should consider the factors such as transaction risks brought by data suppliers, reduced transaction costs and hidden value created. Zhang et al. [62] proposed that the equal revenue sharing mechanism and individual revenue sharing mechanism were more inclined to small data providers, while the marginal utility sharing mechanism was more inclined to big data providers. From the perspective of challenges brought by data economy to government statistics, Xu [63] discussed the contribution degree and reward mechanism of data products and data elements included in the scope of statistical production.

D. FROM THE PERSPECTIVE OF DATA PRODUCT QUALITY MANAGEMENT AND EVALUATION SELECTION

Li and Yao [64] checked and evaluated the accuracy of the digital elevation model data set produced by the Space Shuttle Radar Topographic Mapping Mission data, and analyzed its influence on the accuracy and data quality of the global data set. Song and Qin [4] posited that there are four roles in the manufacturing environment of data products including data provider, data producer, data consumer and data manager. The data quality required four steps of application view establishment, quality parameter determination, quality indicator determination, and quality view integration and application view refinement. Zhang et al. [65] applied the fuzzy gravity center comprehensive evaluation method to assess the quality of GIS data products from five aspects: mathematical accuracy, attribute accuracy, data integrity and accuracy, logical consistency and attachment quality. Wang et al. [66] used the average deviation, standard deviation and correlation coefficient as the indexes of salinity data product quality evaluation. In order to solve the problem of data product selection in hesitant fuzzy language information environment, Chen [67] constructed a data product selection method based on hesitant language judgment matrix, and designed a convergence algorithm to improve the consistency of hesitant language judgment matrix. Dai [68] designed the optimal degree calculation formula of hesitant fuzzy language elements, and constructed the evaluation indexes of data product service provider selection related to product quality, processing ability, storage performance and after-sales service based on hesitant language decision model. Liang [69] combined probabilistic hesitant fuzzy elements and Frank operation to construct a probabilistic hesitant fuzzy information integration algorithm for selecting data products from the dimensions of product quality, product computing speed, reputation and after-sales service. Lin [70] developed a convergent data product selection model based on the language preference relationship consistency improvement algorithm with the language preference relationship information provided by experts. Shi [71] proposed to construct a single-value medium-intelligence multi-attribute decision model with completely unknown attribute weight based on cost-benefit indicators such as product quality, processing ability, purchase cost and after-sales service for the selection of data product service providers. Luo et al. [72] identified the influencing factors of selecting satellite image data products and applied hybrid evaluation information to represent them, and calculated their weight values by some penalty functions and the improved artificial fish swarm algorithm. Huang et al. [2] systematically researched and sorted out the quality assessment indicators of data products (standardization, integrity, accuracy, consistency, timeliness and accessibility), participants (data suppliers, data providers, data exchanges, data demanders and evaluation service institutions), stage product division (original data set, desensitization data set, model data and AI data) and Evaluation methods (computer automatic inspection, computer aided inspection and manual inspection).

In summary, Chinese academia has conducted in-depth discussions on the legal attributes and right-interest ownership of data products, and scholars generally agrees that data products have clear property rights, including exclusivity, possession, use, profit and disposal rights. The research on the content identification and development design of data products involves the classification, development process and technical design of data products, which included the identification of different types of data products, data product generation links, the concept and content level of data services, as well as the market and application platform of data products. Additionally, it discussed the difficulty of data product pricing, different pricing models and revenue distribution mechanisms, as well as different quality evaluation indicators and selection decision models. The research process of data products in Chinese academia reflects the process from preliminary exploration to in-depth research. Early researches mainly focused on the definition and basic theory construction of data products, while recent researches have paid more attention to the practical application fields such as legal attributes, market value, quality management and technology development of data products. Although significant progress has been made in the research of data products, there are still some academic gaps for further exploration, such as: (1) Interdisciplinary research, data products involve law, economics, technology and other fields, interdisciplinary comprehensive research still needs to be further developed; (2) International comparative research, under the background of globalization, comparing the development models, policies and regulations of data products in different countries can provide more references for the development of data products in China; (3) Data security and privacy protection, with the wide application of data products, data security and privacy protection have become an increasingly important issue; (4) The combination of AI and data products, it is necessary to explore how AI technology can promote data more effectively; (5) Value co-creation and revenue distribution of data products, focusing on the role and contribution of multiple stakeholders in the process of value co-creation, and establishing a fair and effective revenue distribution mechanism.

III. RESEARCH METHOD AND DATA SOURCES

CiteSpace, a literature research tool, can visualize complex research fields and reveal the evolution of research topics. It has been widely used in various research fields, such as bibliometric research on information security [73] and social responsibility [74]. The main input data for this paper is drawn from relevant research literature sourced from both CNKI and WOS. CNKI provides Chinese core and above research literature, while WOS offers SCIE and SSCI research literature.

A. RESEARCH METHOD

Based on the foundation of previous research, this paper uses CiteSpace (6.2.R2) software to conduct basic data analysis and visual presentation of literature related to data products

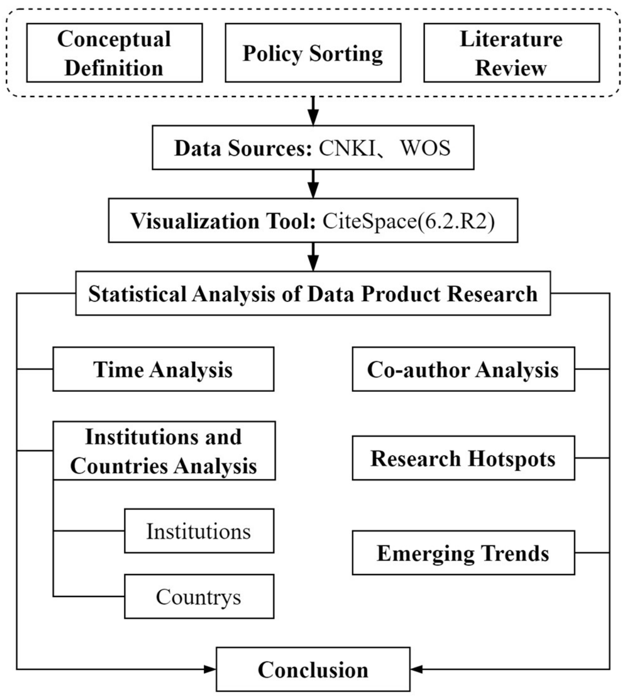

FIGURE 1. Research framework.

in CNKI and WOS, analyzes the theme evolution process and research development trend, and helps scholars systematically grasp the status quo and theme evolution of data product-related research. Figure 1 shows the research framework and main analysis content of this paper. This paper takes the research literature on data products in Chinese academia as the main research object, firstly collects and defines the concept of data products, then sorts out the closely related data element policies, and forms a literature review of the research on data products in Chinese academia. In order to deeply understand the research trends and hot frontiers of data products in academia, based on CNKI and WOS data, this paper uses CiteSpace to conduct econometric analysis of data product research literature from five aspects: publication year, co-authors, institutions and countries, research hotspots and frontier trends. It is aims to provide useful references and suggestions for further research in the field of data products.

B. DATA SOURCES

Based on the research literature data of CNKI and WOS (the year deadline is 2023), this paper constructs a basic analysis database to compare the research content and progress of data products in Chinese and international academia. (1) The Chinese core and above research literature in CNKI are used as the basic data sources, and the advanced search function is used to set the search conditions as the title or keyword containing data products, and the exact matching is adopted by default. After the preliminary search, 143 authoritative articles are obtained. In order to reduce the influence of the research articles with low relevance, this paper eliminates the articles with low relevance to the topic by screening one by one, and finally obtains 137 effective articles. (2) The core collections (SCIE and SSCI) are selected from the official website of WOS. The search criteria are (TI=(data NEAR0 product)) AND TI=(data SAME product), that is, the title contained data product, there is no letter interval between the words data and product, and the types of papers are paper and review paper. Due to the large number of articles retrieved, 399 valid articles were obtained after screening the highly relevant articles.

IV. STATISTICAL ANALYSIS OF DATA PRODUCT RESEARCH

A. TIME ANALYSIS

Figure 2 illustrates the emergence of international academic research on data products starting around 1982, largely driven by advancements in information technology and the encouragement of interdisciplinary collaboration during that period. With the maturity of database technology, the popularization of the Internet, and the rapid development of big data, cloud computing, AI and other technologies, data products are widely used in business decision-making, scientific research, social management and other fields. It leads to an increasing number of scholars and research institutions engaging in the study of data products. In contrast, the research on data products in Chinese academia started relatively late, which is related to the fact that China's information technology and Internet development were still in the initial stage in the early 1990s. However, it has also shown a rapid growth trend since 1996, which is closely related to the rapid development of China's economy and the wide application of information technology. As China has emerged as one of the largest Internet markets globally, the abundance of data resources and advancements in data processing technology have led to a growing demand for data products in the country. In particular, there has been a notable increase in research on data products in recent years. The Chinese government has shown significant commitment to the development of the big data industry by implementing various policies and measures, which have provided substantial support for the research and utilization of data products. During the 32 years from 1982 to 2023, the number of relevant articles retrieved by CNKI and WOS was 137 and 398, respectively. From the change number of publications in each year, the research on data products in China and the world showed an overall upward trend, with a rapid growth trend after 2010 and a historical peak in 2022. The continuous development of global information technology and the increasing richness of data resources have made data products a prominent focus in academia and industry. This trend suggests that data products will have significant implications for future business decision-making, scientific research, social management, and other fields, indicating broad research and development prospects.

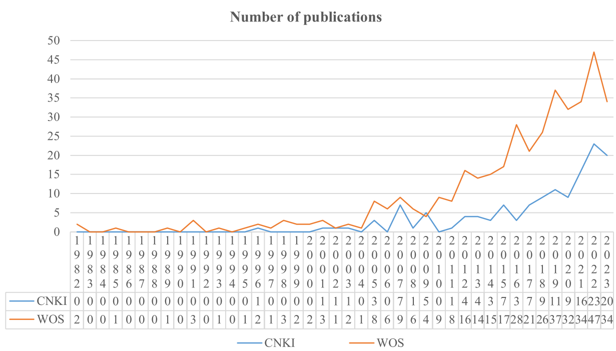

FIGURE 2. Number of publications on data product research.

B. CO-AUTHOR ANALYSIS

This paper takes "co-author" as the node type, and runs CiteSpace after selecting the time span according to the publication time of the literature. In Figure 3, the running results show that the density of author co-occurrence knowledge map in CNKI database for data product research is 0.0102, with 304 nodes and 470 lines. It indicates that there is a certain amount of cooperative research and communication among authors in data product research, forming a relatively close academic community. From the center to the outer circle in Figure 3, the author distribution is from near to far, and there are few cross-team connections, which indicates that there is a lack of cross-team cooperation in this research field. The size of the node represents the published number of authors, the rings in different colors of the same node represent the different publication years of an author, and the darker color means the closer year of publication. Among them, Jin Wang, Jie Zhang and Ronggao Liu et al. 's team maintained a continuous research state in the field of data products.

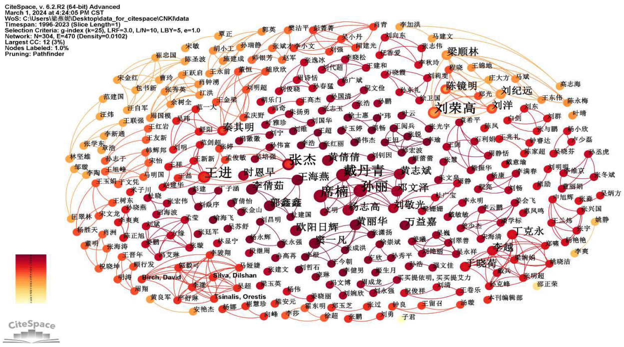

FIGURE 3. CNKI author co-occurrence knowledge map.

Based on the relevant literature samples of WOS database, it shows 632 nodes, 1381 connections and a density of 0.0069 in the WOS author co-occurrence knowledge map of Figure 4. It can be observed that the whole network structure is relatively close, and there are many mutual connections between each author node. It shows that there is extensive cooperation and communication among the authors in the field of data product research in WOS database. Most of the authors tend to research cooperatively, while a few authors prefer to research independently. Chuanmin Hu, David R Doelling and Jean-Christopher Lambert, and other authors showed multiple colors of node rings, indicating their continuous research contributions in this field.

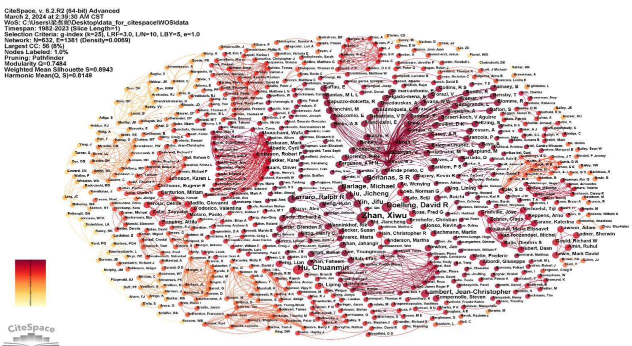

FIGURE 4. WOS author co-occurrence knowledge map.

As shown in Table 3, 137 CNKI literature samples selected in this paper involved 304 authors, and 6 scholars published more than 2 papers in CNKI database. A total of 632 authors were involved in 398 relevant research papers published in WOS database, and 9 scholars published more than 2 papers. According to the data comparison in CNKI database, Ronggao Liu began continuous research on data products in 2007, while other core authors started after 2014. In WOS database, scholars who published more papers on data products have been conducting continuous research since 2016. This suggests that a significant number of core authors in the field of data products have emerged in recent years, with increasing research activity.

TABLE 3. Domestic and foreign authors with more than two publications.

Database	Academic	Earliest Year of Publications	Number of Publications
CNKI	Jin Wang	2014	3
	Danqing Dai	2021	3
	Jie Zhang	2014	3
	Nan Xi	2021	3
	Li Sun	2021	3
	Ronggao Liu	2007	3
WOS	Xiwu Zhan	2016	5
	Chuanmin Hu	2016	4
	David R Doelling	2018	4
	Jicheng Liu	2023	3
	Ralph R Ferraro	2023	3
	Michael Barlage	2023	3
	Jifu Yin	2023	3
	Jean-Christopher Lambert	2018	3
	S R Berlanas	2022	3

C. INSTITUTIONS AND COUNTRIES ANALYSIS

1) INSTITUTION ANALYSIS

The node type is set as "institution," and the research institution co-occurrence knowledge map of CNKI shows a density of 0.0083, 193 nodes and 154 connections in Figure 5. It demonstrates that there are 193 research institutions in CNKI in terms of data products, and there are only

FIGURE 5. CNKI research institution co-occurrence knowledge map.

154 links between each research institution, which indicates that there is a certain cooperative relationship among research institutions, but relatively few. For example, Shanghai Jiao Tong University, Shanghai University of International Business and Economics, and Jiangsu Vocational and Technical College of Food and Drug are identified as conducting independent research. The research of data products has garnered significant interest within the academic community. However, there appears to be a lack of willingness for cooperation and collaboration among research institutions, resulting in a relatively low level of sharing and exchange of knowledge and research results. Many institutions, such as the Chinese Academy of Sciences, the University of Chinese Academy of Sciences, the China Earthquake Networks Center and the State Oceanic Administration, have made continuous research contributions in the research of data products. In addition, it is worth to note that the research institutions are mainly universities and research institutes, and there is a lack of joint research with enterprises and governments. It means that the current data product research is still mainly confined to the academic world, and has not yet formed a research mode of industry-university-research-application integration.

It can be seen from Figure 6 that the density of the research institution co-occurrence knowledge map of WOS is 0.0245, and the node number and connecting lines are 358 and 1565 respectively. It means that there are 358 institutions conducting research on data products in WOS database, there are many mutual connections and cooperative research is relatively frequent, and obvious clustering is formed in cross-institution research. It reflects that research institutions have relatively more cooperation and require to share and exchange research resources with each other, mainly universities, research institutes and government departments, but also lack of joint research between research institutions and enterprises. The node circles display multiple colors, such as National Aeronautics & Space Administration, National Oceanic Atmospheric Admin – USA, NASA Goddard Space Flight Center, Chinese Academy of Sciences and Centre National de la Recherche Scientifique. It indicates that these

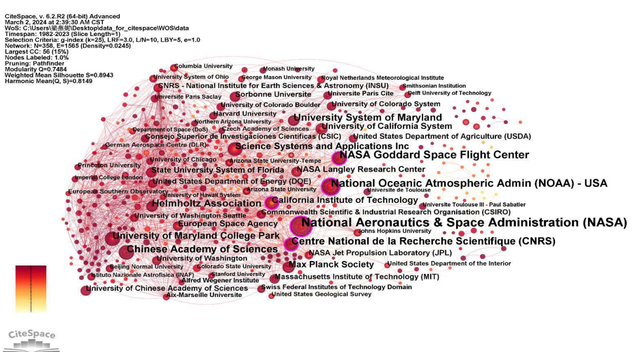

FIGURE 6. WOS research institution co-occurrence knowledge map.

institutions have carried out continuous exploration in the field of data products.

TABLE 4. Domestic and foreign institutions with a large number of publications.

Database	Institution Name	Earliest Years of Publications	Number of Publications
CNKI	China Earthquake Network Center	2019	5
	Institute of Geographic Sciences and Resources, Chinese Academy of Sciences	2005	5
	University of Chinese Academy of Sciences	2014	5
	Institute of Remote Sensing Applications, Chinese Academy of Sciences	2007	3
	Graduate School of the Chinese Academy of Sciences	2009	3
WOS	National Aeronautics & Space Administration	1991	76
	National Oceanic Atmospheric Admin - USA	2003	40
	NASA Goddard Space Flight Center	1998	40
	Chinese Academy of Sciences	2012	33
	Centre National de la Recherche Scientifique	2005	31
	Helmholtz Association	2005	27
	University System of Maryland	2001	26
	University of Maryland College Park	2003	23
	Science Systems and Applications Inc	2001	22

According to the statistics in Table 4, in CNKI database, China Earthquake Networks Center, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences and University of Chinese Academy of Sciences all published 5 papers, which are the highest research institutions. And the Institute of Remote Sensing Applications of the Chinese Academy of Sciences and the Graduate School of the Chinese Academy of Sciences followed with 3 publications each. In WOS database, National Aeronautics & Space Administration published 76 papers, which is the research institution with the highest number of papers. National Oceanic Atmospheric Admin - USA and NASA Goddard Space Flight Center followed with 40 publications each. It shows that universities and research institutions play a crucial role in the field of data product research, both in China and internationally. These institutions have professional research teams, rich research resources and advanced experimental equipment, which provide strong support for the research in the field of data products. In addition, the active participation of universities and research institutions at home and abroad has also promoted the research progress and technological innovation in the field of data products. Through in-depth exploration and practice, they have provided rich experience and achievements for the development, application and optimization of data products, which has promoted the rapid development of data products.

2) COUNTRY ANALYSIS

It can be seen in Figure 7, the density of the research institution co-occurrence knowledge map of WOS is 0.2417, and the node number and connecting lines are 62 and 457 respectively. It means there are 62 institutions conducting research on data products in WOS database. The high number of mutual connections indicates a significant level of transnational collaborative research. This highlights the global focus on data product research and the increasing trend of cooperation. Additionally, it suggests the presence of clear research strengths and popular topics in this field, emphasizing the significance of international collaboration. The United States, China, Germany, France and other countries show multiple colors of node rings, indicating that these countries have carried out continuous exploration in the field of data products.

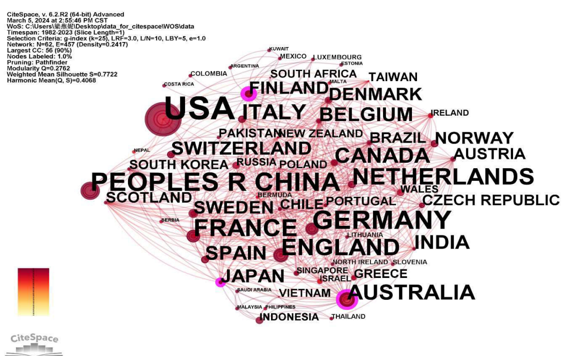

FIGURE 7. WOS country co-occurrence knowledge map.

With the rapid development of digitalization, informatization and big data technology, data products have become the focus of research worldwide. As shown in Table 5, on the data product research of the WOS database, the United States has published the highest number of 227 papers since 1991, far exceeding other countries. It is closely related to the United States' advantages in scientific and technological innovation, R&D investment and policy support. China and Germany followed with 68 and 67 papers, respectively. It demonstrates that China and Germany are increasing their investment in the field of data products, and strive to occupy a favorable position in the global scientific research competition. In addition, France, England, the Netherlands, Australia, Canada, Italy, Spain, Belgium and India all published more than 20 related papers. Among them, the publication years of India and Canada are the earliest, which are 1982 and 1985, respectively. It is related to the increased attention and investment in the field of data products in these countries.

TABLE 5. Countries with more than 20 publications.

Database	Country	Earliest Year of Publications	Number of Publications
WOS	USA	1991	227
	PEOPLES R CHINA	2007	68
	GERMANY	2005	67
	FRANCE	2004	49
	ENGLAND	2010	40
	NETHERLANDS	2001	40
	AUSTRALIA	2011	32
	CANADA	1985	32
	ITALY	1999	29
	SPAIN	2010	23
	BELGIUM	2005	21
	INDIA	1982	21

Table 6 and Figure 8 cluster the literature data of the WOS database with countries as node keywords with CiteSpace. The information of the country node keyword co-occurrence network cluster table shown in Table 6 is obtained through the menu bar of ''Summary Table | Whitelists''. From this, it can be judged the hot topic words those under the same keyword cluster in the field of data product research in recent years. The cluster average contour values are all greater than 0.7, indicating that the cluster is efficient and convincing. The average year is the average year when the hot words firstly appeared in each cluster. Cluster 0 mainly focuses on soil moisture and sea surface observation techniques, especially research related to SMAP tasks. At the same time, it also involves active and passive absorption line techniques related to remote sensing observations, with the average year of 2010. Cluster 1 mainly focuses on techniques for modeling and calibration in data constrained areas, particularly using Bayesian calibration and ensemble Kalman filtering methods, as well as the application of process-based GRASS models and data products, with the average year of 2012. Cluster 2 focuses on the natural water balance in Southeast Asia and the impact of deforestation on the environment, particularly research related to mangrove sustainability, as well as general themes of galaxy research, with the average year of 2013. Cluster 3 mainly focuses on observation methods and cosmological observations, especially topics related to galaxy and stellar research, as well as large-scale astronomical survey projects and data analysis methods, with the average year of 2015. Figure 8 generates a country time zone map clustered by keywords with selecting the ''layout | Timezone View'' option in CiteSpace. Under the same keyword

TABLE 6. Country node keyword co-occurrence network cluster table.

Cluster Number	Cluster Size	Cluster Average Contour Value	Average Year	Identifier Word (LSI)	Research Country
0	21	0.742	2010	soil moisture; sea surface; smap mission; core validation sites; active passive | absorption lines; english channel; epos delivery framework; magnetospheric dynamics; reference frames	Canada; USA; Italy; Netherlands; France; Finland; Germany; Peoples R China; Switzerland; England; Nepal; Pakistan; South Africa; Malta; Luxembourg; Colombla; Argentina; Costa rica; Mexico; Estonia
1	14	0.778	2012	data-limited areas; bayesian calibration; ensemble kalman filter; process-based grass models; data products | stellar content; process-based grass models; data products; open clusters; data-limited areas	Taiwan; South Korea; Norway; Israel; Spain; Scotland; Ireland; New Zealand; Portugal; Brazil; Wales; Russia; Singapore
2	11	0.782	2013	southeast asia; natural water balance; deforestation depletion; mangrove sustainability; galaxy: general | soil moisture; active passive; core validation sites; satellite broadcasting; soil measurements	India; Sweden; Australia; Vietnam; Japan; Serbia; Indonesia; Philippines; Thailand; Malaysia; Saudi arabia
3	10	0.817	2015	methods: observational; galaxy: general; stars: general; surveys; cosmology: observations | magnetic field; comprehensive inversion; electromagnetic induction; swarm satellites; stars: general	Belgium; Austria; Poland; Czech Republic; Denmark; Chile; Greece; Slovenia; North Ireland; Lithuania

Note: The country attributes are summarized based on the country time zone map information of the data product in Figure 8.

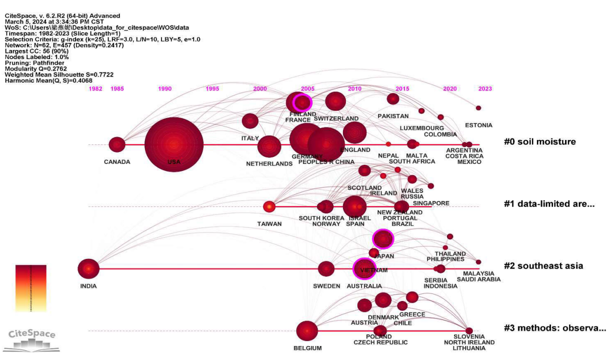

FIGURE 8. WOS data product research national time zone map (LSI).

cluster, different countries will be presented on the timeline based on their first appearance year. The line indicates that two countries appear in the same or multiple papers, and the higher the frequency of country appearance, the larger the circle. Under the theme of soil moisture and sea surface observation in cluster 0, the main research countries in chronological order include Canada, USA, Italy, Netherlands, France, Finland, Germany, Peoples R China, Switzerland, England, Nepal, Pakistan, South Africa, Malta, Luxembourg, Colombla, Argentina, Costa rica, Mexico, Estonia and so on. Under the theme of Data Modeling and Calibration in Cluster 1, the main research countries in chronological order include Taiwan, South Korea, Norway, Israel, Spain, Scotland, Ireland, New Zealand, Portugal, Brazil, Wales, Russia, Singapore and so on. Under the theme of Southeast Asian Environment and Galaxy Research in Cluster 2, the main research countries in chronological order include India, Sweden, Australia, Vietnam, Japan, Serbia, Indonesia, Philippines, Thailand, Malaysia, Saudi arabia and so on. Under the theme of Cosmological Observations and Stellar Research in Cluster 3, the main research countries in chronological order include Belgium, Austria, Poland, Czech Republic, Denmark, Chile, Greece, Slovenia, North Ireland, Lithuania and so on. The above also demonstrates the national attributes of different research topics in the field of data product research in the international academic community.

D. RESEARCH HOTSPOTS

The keyword clustering knowledge graph is generated for further analysis, and the tags with obvious clustering are selected in Figure 9. These clusters reflect the hot issues in the domestic and international research of data products. CNKI includes nine cluster labels: data product, data pricing, data resources, big data transaction pricing, drought index, network operator, big data product, Yangbi earthquake in Yunnan province and hesitant language judgment matrix. The emergence of three cluster labels—data product, data pricing, and data resource—highlights the importance of the data product itself and its value assessment and resource management in current research. It reflects that, in the digital era, how to effectively develop, utilize, and price data products has become a key issue. As a separate cluster label, big data transaction pricing shows that the pricing mechanism and strategy of the big data transaction market is also a hot issue attracting much attention. With the rapid development of big data industry, how to formulate reasonable transaction pricing strategy is both a challenge and an opportunity. Drought index, as a cluster label, reflects that in the application field of data products, the cross-research of environment and climate has also become a hot topic. It indicates that data products are not only widely used in business and technology fields, but

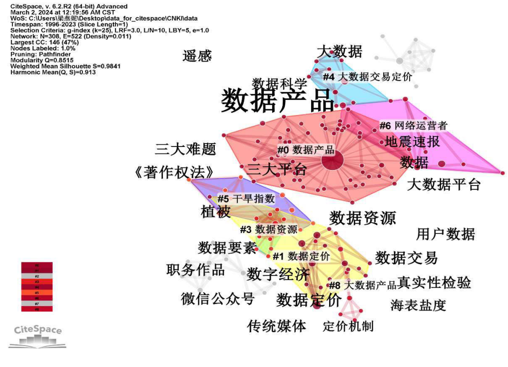

FIGURE 9. CNKI keyword clustering knowledge graph (LSI).

also play an important role in environmental protection and disaster management. Additionally, network operator, as a clustering label, reveals the important role of data products in network operation, which is closely related to network security, data privacy protection and other related issues. As a cluster label, big data product involves the development, optimization and promotion of big data products, which reflects the position and influence of big data products in the market. The two cluster labels of Yunnan Yangbi earthquake and Hesitant language judgment matrix are related to specific research events or methods, reflecting the application and discussion of data products in specific events or specific research methods.

WOS includes soil moisture, backscatter ultraviolet, precipitation, data products, CO_2, ocean color, absorption, atmosphere and random forest and other 14 cluster labels. Figure 10 shows the keyword clustering knowledge map displayed by setting CiteSpace's "Show the Largest K Clusters" to 9. The soil moisture clustering label involves soil moisture data sets, model and algorithm data products, as well as soil moisture monitoring, measurement of impacts on climate, agriculture and ecosystems. The backscatter ultraviolet clustering label displays the application research of ultraviolet backscatter measurement data and its data products in climate monitoring, air quality assessment, and other fields. The precipitation clustering label involves data products such as precipitation data sets, precipitation models and forecast systems, and their impacts on climate, hydrology and ecology. The data products clustering label is directly involved in the development, application and evaluation of data products, including various types of data sets, algorithms, models and software tools, which are used to support research in environmental, ecological, climate and other fields. The CO_2 clustering label is concerned with data products such as carbon dioxide concentration data sets, emission inventories and prediction models, and their impacts on climate change. The ocean color clustering label involves data products such as ocean color remote sensing data sets, image processing results and analysis reports, as well as its applications in Marine ecology, Marine environment and fishery. The absorption clustering label involves the absorption spectrum data set, absorption coefficient measurement results, absorption process data and model research of substances in different media, such as water, atmosphere, soil, etc. The atmosphere clustering label involves data products such as atmospheric composition data sets, meteorological observation data and climate model output, and its applications in climate change, air quality and other aspects. The random forest clustering label refers to the use of random forest algorithm for data analysis and modeling, or the use of data analysis tools and models developed by random forest algorithm. Based on the above, there are abundant researches on the concept and interests of data products in Chinese academic circles. Moreover, the comprehensive research on the transaction and operation of data products is more prevalent on this basis. In WOS database, most relevant literature

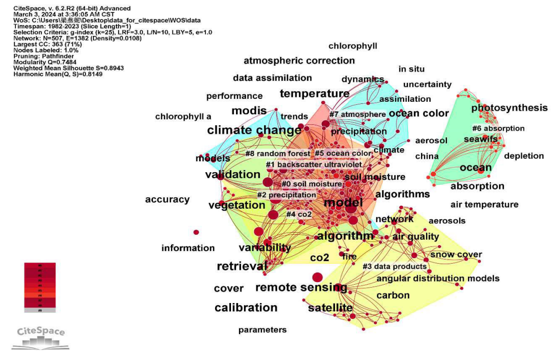

FIGURE 10. WOS keyword clustering knowledge graph (LLR).

obtained using data product as a title search term is about the development and application of data products in specific fields, such as astronomy and geography. It shows that Chinese academia and international academia have different focuses and advantages in data product research, but also potential for mutual complementation and cooperation.

The CNKI keyword clustering knowledge graph outputs the information of different keywords, such as the frequency, centrality and the earliest occurrence year in Table 7. The data product, as the central word, had the highest frequency and centrality of 54 times and 0.42 respectively, followed by data pricing, data transaction and data elements, which

TABLE 7. Analysis of frequent keywords.

Database	Frequency	Center Degree	Year	Keyword
CNKI	54	0.42	2007	data product
	6	0.02	2018	data pricing
	5	0.03	2018	data trading
	5	0.02	2022	data element
	4	0.03	2021	big data
	4	0.08	2021	digital economy
	4	0	2019	earthquake quick report
	4	0	2007	remote sensing
	3	0.04	2005	vegetation
	3	0.01	2022	pricing mechanism
	3	0	2022	enterprise data
	3	0.02	2005	data
	3	0.03	2016	data resource
WOS	29	0.08	2014	validation
	27	0.16	2005	model
	20	0.07	2013	variability
	19	0.1	1998	remote sensing
	19	0.15	2009	climate change
	17	0.1	2000	algorithm
	17	0.14	2006	retrieval
	15	0.09	1991	satellite
	13	0.08	2010	modis
	13	0.06	2011	vegetation
	13	0.09	2006	temperature
	12	0.12	1993	calibration
	12	0.02	2018	soil moisture
	12	0.02	2019	precipitation

Note: The keyword frequency of CNKI is more than 2 times, and the keyword frequency of WOS is more than 10 times.

were related to data definition and operation. Big data, digital economy, quick earthquake reporting, remote sensing and vegetation followed closely, showing the specific application fields of data products. The most frequent keywords of data product related research in WOS database were validation, model, variability, remote sensing, climate change, algorithm, retrieval, satellite, modis, vegetation, temperature, calibration, soil moisture and precipitation. It demonstrates that these keywords occupy the core position in the international research related to data products. Among them, model, climate change, retrieval and calibration have the highest centrality, which are 0.16, 0.15, 0.14 and 0.12, respectively, indicating the application and research focus of data products in remote sensing, meteorology and environmental science.

Select the ''Summary Table | Whitelists'' in the ''Clustern'' menu bar, which can obtain the keyword co-occurrence network clustering table information of Table 8. It can determine the hot topic keywords in the field of data product research in recent years. If the cluster average contour value is more than 0.7, it represents the clustering is efficiency convincing. The average year is the year in which the hot topics first appear in each cluster, and the result is obtained by averaging the years. In recent years, the research hotspots of data products and related fields of CNKI are mainly reflected in data pricing, transaction, digital economy, big data assets, remote sensing observation, personal information protection, seismic monitoring and decision-making methods. Cluster 0 focuses on the rights allocation and remote sensing observation of data products, with the average year of 2018, indicating that the application of data products in remote sensing observation is increasingly important. Cluster 1 involves data pricing, transactions and the digital economy, with the average year of 2021, reflecting the rapid development of the data market and the rise of the digital economy. Cluster 3 focuses on data resource management, copyright law and other aspects, with the average year of 2017, indicating that the legal and management issues of data resources get recognition gradually. In addition, clusters 4 and 5 focus on big data transaction pricing, right confirmation and drought monitoring, with the average years of 2020 and 2008 respectively, indicating that the value of big data and remote sensing applications are current research hotspots. Cluster 6 involves network operators and personal information, with the average year of 2013, reflecting the importance of personal information protection in the Internet era. Clusters 8 and 9 focus on the legal protection of big data products and the application of seismic monitoring data products, with the average years of 2014 and 2020, respectively. The average year of cluster 15 is 2017, which reflects the application trend of data products in decision support system.

The research hotspots of data products and related fields in WOS database are mainly concentrated in soil moisture, atmospheric science, remote sensing, oceanography, astronomy and climate change. Clusters 0 and 1 involved keywords such as soil moisture, data assimilation, nitrogen dioxide and air quality, with the average year of 2015, reflecting the research trend of atmospheric and environmental sciences. Cluster 2 focuses on precipitation observation, data integration and statistical analysis methods, indicating that these fields were the focus of research around 2012. Cluster 3 involves data products, angle distribution model, radiation budget, etc, with the average year of 2011, which reflects the research focus of remote sensing data products. Clusters 4 to 6 are related to oceanography and geochemistry, such as carbon dioxide, ocean color and chlorophyll, with the average years of 2007 and 2005, respectively. Clusters 7 and 8 focus on climate change and the application of machine learning to data products, respectively, which shows the research direction in these areas, with the average years of 2011 and 2017, respectively. Cluster 10 focuses on keywords such as aerosol optical thickness and remote sensing inversion, with the average year of 2012. Cluster 11 involves remote sensing data processing, geospatial data, etc, with the average year of 2008. Cluster 12 researches on satellite images, mesoscale vortices, etc, with the average year of 2012. Cluster 15 focuses on keywords related to atmospheric aerosols such as dust and particle density, and the average year is 2010. Cluster 17 deals with astronomical databases, galaxy evolution, etc, with the average year of 2023. In summary, the research hotspots and development trends of data products and related fields cover a variety of disciplines, including computer science, statistics, physics, environmental science, geography, economics, etc. The application and development trends of data products in different fields are important directions for current and future research. These interdisciplinary research hotspots not only reflect the wide value of data products in practical applications, but also predict the new trend of scientific research and industrial development in the future.

E. EMERGING TRENDS

Research frontiers refer to emerging theories and hot topics in a research field, which can be analyzed and judged by keyword time zone maps and emergent words. The keyword time zone map can be generated by selecting the ''Layout | Timezone View'' option of CiteSpace software. In Figures 11-12, the same clustering keywords in CNKI and WOS can present according to their first year on the time line, the wired means two key words appear in the same paper or multiple papers, and the high keyword frequency the bigger circle. According to the clustering data analysis of CNKI literature, the research on data products has been continuously carried out since 2000, and the co-occurrence relationship is obvious. (1) The keyword timeline of data product co-occurring shows that data products focused on market expansion and service systems in the early stage, and then turned to the fields of meteorology and earth science. With the rise of satellite remote sensing, information services and surface monitoring, data products have become more widely used. (2) The keyword timeline of data pricing co-occurrence indicates that data pricing research has gradually deepened from basic concept to practical application. Value evaluation methods have

TABLE 8. Keyword co-occurrence network clustering table.

Database	Cluster Number	Cluster Size	Cluster Average Contour Value	Average Year	Identifier Word
CNKI	0	60	0.998	2018	data product; right configuration; voluntary registration; frank norm; remote sensing observation | target recognition; remote sensing observation; empirical/semi-empirical model; physical model; parameter extraction
	1	22	0.988	2021	data pricing; data trading; digital economy; data market; data elements | data assets; data bound; public data; price calculator; digital economy
	3	12	0.958	2017	data resource; traditional media; the copyright law; data product; Job work; big data platform; user data; three major platform; WeChat official account; three difficult problems
	4	11	0.935	2020	big data transaction pricing; big data asset; big data ownership; data science; archive work | file management; data science; archival work; big data transaction pricing; big data asset
	5	11	0.988	2008	drought index; drought monitoring; apparent thermal inertia; modis; data product | data product; remote sensing inversion; leaf area index; aridity index; drought monitoring
	6	10	0.972	2013	network operator; emerging rights; data products; personal information; accurate personal information | network operators; the network user information; emerging rights; data products; personal information
	8	8	0.985	2014	big data product; property of right; property ownership; legal protection; and anti-unfair competition
	9	7	0.985	2020	Yunnan yangbi earthquake; earthquake quick report; aftershock; focal mechanism; earthquake intensity; data product
	15	5	0.99	2017	hesitant language judgment matrix; additive consistency; consistency index; decision making method; data product
WOS	0	45	0.811	2015	soil moisture (18.15, 1.0E-4); data assimilation (11.73, 0.001); nitrogen dioxide (7.81, 0.01); air quality (7.81, 0.01); boundary layer (7.81, 0.01)
	1	42	0.833	2015	backscatter ultraviolet (8.27, 0.005); air (8.27, 0.005); assimilation (8.27, 0.005); aerosol (8.27, 0.005); drought (8.27, 0.005)
	2	40	0.88	2012	precipitation (13.32, 0.001); accuracy (4.5, 0.05); nam co (4.5, 0.05); asian precipitation highly resolved observational data integration toward evaluation of the water resources (aphrodite) (4.5, 0.05); markov chain monte carlo (4.5, 0.05)
	3	38	0.884	2011	data products (14.36, 0.001); angular distribution models (14.36, 0.001); radiation budget (14.36, 0.001); methodology (14.36, 0.001); clouds (9.55, 0.005)
	4	30	0.927	2007	co2 (18.39, 1.0E-4); potential temperature (13.76, 0.001); V (13.76, 0.001); alkalinity (13.76, 0.001); water (13.76, 0.001)
	5	26	0.891	2009	ocean color (22.38, 1.0E-4); seawifs (16.19, 1.0E-4); satellite validation (11.12, 0.001); chlorophyll (11.12, 0.001); water-leaving radiance (11.12, 0.001)
	6	24	0.964	2005	absorption (11.44, 0.001); photosynthesis (11.44, 0.001); ocean (6.13, 0.05); color data products (5.7, 0.05); oceanic waters (5.7, 0.05)
	7	24	0.972	2011	atmosphere (11.27, 0.001); temperature (7.89, 0.005); arctic ocean (5.62, 0.05); groundwater sustainability (5.62, 0.05); latent heat (5.62, 0.05)
	8	24	0.919	2017	random forest (10.61, 0.005); variability (5.36, 0.05); precipitation measurement (5.29, 0.05); boruta (5.29, 0.05); atlantic (5.29, 0.05)
	10	21	0.937	2012	aerosol optical thickness (11.27, 0.001); retrieval (10.48, 0.005); aeronet oc (7.58, 0.01); reflectance (5.97, 0.05); sounders 3i (5.62, 0.05)

TABLE 8. ***(Continued.)*** **Keyword co-occurrence network clustering table.**

Data base	Cluster Number	Cluster Size	Cluster Average Contour Value	Average Year	Identifier Word
	11	19	0.874	2008	remote sensing (16.42, 1.0E-4); data processing (12.7, 0.001); geospatial data (6.32, 0.05); nohrsc (6.32, 0.05); remap (6.32, 0.05)
	12	17	0.932	2012	satellite imagery (6.69, 0.01); mesoscale eddies (6.69, 0.01); census (6.69, 0.01); information (6.69, 0.01); multi-angle imaging (6.69, 0.01)
	15	7	0.981	2010	dust (11.37, 0.001); particle (7.52, 0.01); density (7.52, 0.01); atmospheric aerosol (7.52, 0.01); sage ii (7.52, 0.01)
	17	6	0.991	2023	astronomical data bases: catalogues (18.61, 1.0E-4); astronomical data bases: atlases (9.18, 0.005); galaxies: evolution (9.18, 0.005); stars: fundamental parameters (9.18, 0.005); galaxy: stellar content (9.18, 0.005)

been explored in the early stage. With the development of data industry increasingly emphasize on data property rights protection and capitalization, scholars began to study the role of data in value creation after entering the digital economy era. In the later period, keywords such as auction design and government statistics reflect the diversity of data pricing practices and policy orientation. (3) The keyword timeline of three platforms co-occurrence shows that keywords focus on the ownership and utilization of data resources from the legal basis of the Copyright Law, and then discuss the relationship between the state ownership of data resources and the market formation mechanism. Finally, they focus on the data utilization and emphasize the effective allocation and utilization of data resources under the premise of legal compliance. (4) The keyword timeline of big data co-occurrence indicates that scholars paid attention to data processing and analysis methods in the early stage, and later focus on data-driven scientific discovery and decision support. With the emergence of keywords such as data tax and archive work, it reflects the application and challenge of big data in social economy and cultural fields. (5) The keyword timeline of vegetation co-occurrence shows that scholars explored the relationship between vegetation and climate in the early stage, and then focused on the monitoring of vegetation by remote sensing technology, and finally researched the close relationship between vegetation and ecological environment, agricultural production, as well as the regional characteristics and practical applications of vegetation research. (6) The keyword

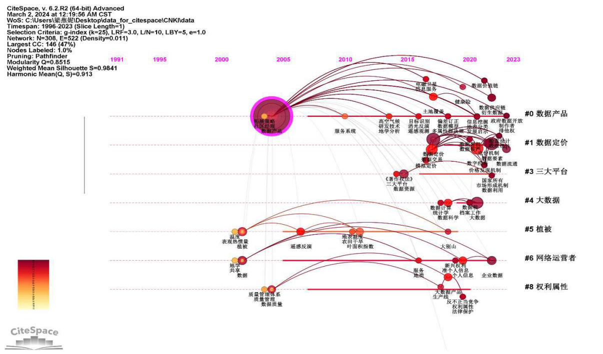

FIGURE 11. **CNKI data product research keyword time zone map.**

timeline of network operators co-occurrence demonstrates that scholars started from basic fields such as geoscience and geology, and gradually paid attention to the application of sharing and data, and then paid attention to the content of data service, data management and user privacy protection. (7) The keyword timeline of rights attributes co-occurrence indicates scholars pay attention to the basis of data quality, and then explored the importance of data quality in the era of big data, the relationship between data quality and business process compliance, and the legal balance between data utilization and protection.

Through comparison, it can be seen that the research on data products of WOS literature is relatively rich, with more effective clusters generated, complete research timelines, and obvious co-occurrence relationship, which reflects the great value of data products in soil moisture, backscattered ultraviolet light, precipitation, remote sensing technology, earth science and climate change, carbon emissions, Marine ecology and other research fields. It also reveals that the research in this field is gradually becoming in-depth and diversified. (1) The keyword timeline of soil moisture co-occurrence shows that soil moisture research is closely related to data products. Early algorithm development lays the foundation for soil moisture monitoring, and then satellite remote sensing technologies such as MODIS and ASTER provide a large number of continuous data for research. As scholars pay attention to the interaction between soil moisture, ecosystems, and air quality, data verification and performance evaluation have highlighted the importance. In recent years, satellite observations and remote sensing reflectance data have provided a new perspective on soil moisture. (2) The keyword timeline of backscatter ultraviolet co-occurrence shows that backscatter ultraviolet research is closely related to data products, which were initially used to study surface energy balance and latent heat flux, and later extended to construct and verify surface reflection models, and then to study atmospheric aerosols and air quality. In addition, data products are also widely used in climate change adaptation and trend analysis, involving ecosystems, hydrology and other fields in many countries and regions around the world. (3) The keyword timeline of precipitation co-occurrence indicates that the data product is first used to describe the distribution and parameters of precipitation, and then explore the relationship between vegetation cover and precipitation, and evaluate the impact of climate change on precipitation. Meanwhile, remote sensing data such as brightness temperature information and radar data play a key role in precipitation monitoring. In addition, data products also support the construction of decision support system and model for flood prevention and control, and are widely used in the fields of bioclimatology and ecology to improve the accuracy and reliability of precipitation research through advanced statistical methods. (4) The keyword timeline of data products co-occurrence shows the in-depth exploration of data products in satellite remote sensing technology, temperature detection technology, description of geophysical processes, climate change, cloud detection, radiation budget and aerosol, and emphasizes the research of data product extraction technology, data fusion and evolution. (5) The keyword timeline of CO_2co-occurrence indicates that early AVHRR and other remote sensing data were used for monitoring and calibration of CO_2. As the accuracy and calibration of the data products became critical, statistical methods such as hypergeometric distribution and time series analysis were applied to the data to reveal the variation trend of CO_2 concentration and the impact of anthropogenic emissions. Additionlly, ocean data products are also crucial for the study of ocean acidification and carbon cycle. (6) The keyword timeline of ocean color co-occurrence shows rich data products related to ocean color in coastal and specific sea areas, involving global chlorophyll monitoring, optical property analysis and other applications, which not only supports the development of relevant algorithms, but also provides support for the study of climate change and Marine ecosystem. In addition, remote sensing instruments and atmospheric correction methods also play an important role. (7) The keyword timeline of absorption co-occurrence shows that the data products of absorption research involve ocean, atmosphere and other aspects, from the early use in ocean water color observation, gradually expanded to study the relationship between phytoplankton fluorescence and absorption, as well as the absorption process of ozone in the atmosphere. In addition, the data products also support fine analysis of absorption spectra, evaluation of absorption efficiency, and identification of different types of absorbers. (8) The keyword timeline of atmosphere co-occurrence demonstrates that the data products of atmospheric research involve climate, temperature, ocean and other aspects, which are widely used in climate monitoring, ocean-atmosphere interaction research, topography and space detection technology. In addition, data products are also used in remote sensing instrument calibration and space detection, environmental protection and climate prediction. (9) The keyword timeline of random forest co-occurrence shows the combination of data products and random forest is used to monitor environmental ecology, analyze remote sensing data, evaluate pollution degree, explore ocean and atmospheric flux, predict basin dynamics, observe land use/cover change, etc. Additionally, this combination is also applied to model accuracy evaluation and error analysis to improve the accuracy of prediction and classification.

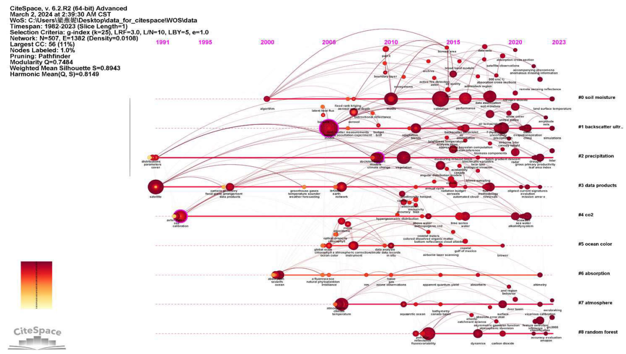

FIGURE 12. WOS data product research keyword time zone map.

Figure 13 shows the node occurrence words with high saliency value in the research literature of data products in CNKI and WOS databases, revealing the research frontier and evolution trend of this field in different time periods. According to the data source of CNKI database, 13 node occurrence words with high occurrence value are obtained. The emergence time of the early frontier is from 2008 to 2013, and the research frontier in this stage mainly focuses on data service and remote sensing inversion, which reflects that researchers began to pay attention to the importance of data service and the application of remote sensing inversion technology in data products during this period. The emergence time of the mid-stage frontier was from 2014 to 2018, and the research frontier in this period mainly focused on authenticity test, sea surface salinity, additive consistency and land cover, which indicated that researchers began to pay attention to the quality and reliability of data products, especially in the field of remote sensing monitoring and geographic information science. As for the latest research frontiers of data products, from 2019 to 2023, the research topics focus on quick earthquake reporting, big data products, data science, personal information, digital economy, data transaction and seismic intensity. It indicates that with the rapid development of big data and digital economy, the application of data products in different fields has become more and more extensive, especially in the fields of earthquake monitoring, personal information protection and digital economy.

Based on the literature samples of data product research in WOS database, 12 node occurrence words with high salience value are obtained in Figure 14. The early frontier emergent time is from 2013 to 2014, the research frontiers of this period mainly concentrated in galaxies: abundances, quasars: absorption lines and algorithm, which indicates that the phase researchers begin to pay close attention to in the field of astronomy data application, especially the important role in the study of algorithm in galaxies and quasars. The significant emergence period of the mid-term frontier spanned from 2016 to 2017. In this stage, scholars mainly focused

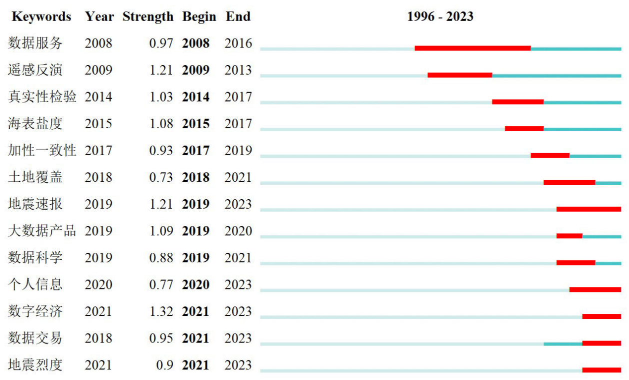

FIGURE 13. CNKI data product research emergent node vocabulary map.

on validation, performance and variability, which reflected researchers' emphasis on the quality and performance evaluation of data products. At the same time, researchers began to explore the variability of data products in different application scenarios. The latest research frontier of data products will be from 2018 to 2023, research topics in this stage include soil moisture, data assimilation, precipitation, climate change, vegetation and temperature. It indicates that the application of data products in the fields of earth science, environmental science and climate research is becoming more and more extensive, which has become a hot topic for current and future research.

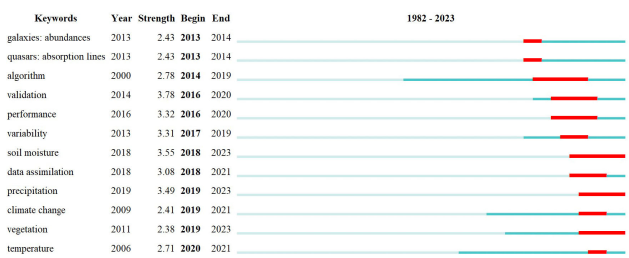

FIGURE 14. WOS data product research emergent node vocabulary map.

In CNKI database, the high emergent value node occurrence words of data product research focus on data service and remote sensing inversion in the early stage, to the quality and reliability of data products in the middle stage, and then to the latest research frontier focusing on the application of seismic monitoring, big data products, digital economy and other fields. In WOS database, the research on data products focuses on the application of data products in astronomy in the early stage, to the quality and performance evaluation of data products in the middle stage, and then to the application of earth science, environmental science and climate research in the latest research frontier. By comparison, it can be found that the emphasis of the two databases in the field of data product research is different. CNKI pays more attention to the application of remote sensing technology and data services, while WOS pays more attention to the data processing and analysis in the field of astronomy. The mid-term research frontiers of the two databases both involve the quality and performance evaluation of data products, which indicates that researchers are beginning to interest in the quality and performance of data products. In addition, both databases reflect the application of data products in geoscience, environmental science and climate research, as well as the focus of big data and digital economy on the opportunities and challenges brought by data product research.

V. CONCLUSION

The research of data products has garnered significant attention in the digital economy. This paper delves into the academic research surrounding data products in Chinese academia, offering a definition of the concept, outlining relevant policies concerning data elements, and presenting a literature review. It conducts thorough econometric analysis on data product research literature from CNKI and WOS databases with CiteSpace, encompassing various dimensions including publication year, co-authors, institutions and countries, research hotspots, and emerging trends. It is aims to comprehensively grasp the dynamics and frontiers of data product research, providing valuable insights and recommendations for future academic and practical applications. Based on previous research, the following conclusions can be drawn. (1) Data products are characterized by multidimension and complexity. The definition of data products in Chinese academia involves data quality, pricing, legal protection, market mechanism, information technology and other aspects. Data products are the main transaction objects of data markets, which are not only the collection of data, but also the products with commercial value and practical application value formed after specific processing and processing. It can be data sets, analysis reports, prediction models, etc. (2) The development of data products is closely related to the policy and legal environment. China pays great attention to data products and their management, market development, standardization, capitalization and legal protection. China has formulated a series of national policies related to data products and data elements in normalizing data management systems, developing data factor markets, promoting data property rights and transaction mechanisms, and standardizing data products and services. It is aims to achieve the standardized management and value realization of data elements. (3) The research of data products is progressing from preliminary exploration to in-depth and multidimensional investigation. Research on data products in Chinese academia has transformed from definitions and basic theory construction to practical applications. However, there are still some research fields that need to be further explored, such as interdisciplinary research, international comparative research, data security and privacy protection, the combination of artificial intelligence and data products, and the value co-creation and income distribution of data products.

According to the analysis results of CiteSpace software, it can be known as following. (1) Data product research has gradually attracted the attention of international academia since 1982. With the rapid development of information technology and the increasing abundance of data resources, its application fields have been expanding. Despite starting later than other countries, China's research output has experienced rapid growth in recent years, driven by policy support and market demand. It is anticipated that data products will increasingly influence business operations, scientific advancements, and social management in the future. (2) A large number of core authors and institutions have emerged in recent years, and the research heat of data products is gradually rising. The research institutions are mainly universities and research institutes. The trend towards international cooperation is clearly visible. Nevertheless, cross-team and cross-institution collaboration within the Chinese academic community is relatively uncommon, and there is a lack of joint research initiatives involving enterprises and governments. (3) The application and research of data products in remote sensing, meteorology, environmental science, and other fields are well represented in CNKI and WOS databases. Chinese academia primarily focuses on comprehensive research, particularly in the areas of concept, rights and interests, transaction, and operation of data products. In contrast, international academia tends to emphasize the development and application of data products in specific fields like astronomy and geography. (4) Research hotspots in data products and related fields widely cover various disciplines such as computer science, environmental science, and geography. These areas include data pricing, trading, remote sensing observation, personal information protection, earthquake monitoring, and more. The focus is on the application, value, law, and management issues of data products, along with research trends in remote sensing, atmospheric science, and climate change. (5) Chinese academic research on data products has been steadily advancing since 2000, initially concentrating on market growth and service systems before branching out into areas such as meteorology and earth science. The research emphasis has evolved from data processing and value assessment to data-driven scientific exploration, decision-making support, and efficient data resource utilization, with a strong emphasis on data quality and legal adherence. (6) The international academic research on data products encompasses various areas such as soil moisture, backscattered ultraviolet light, precipitation, remote sensing technology, earth science, and climate change. The research trend shows a progression from basic to in-depth analysis, and from focusing on single variables to exploring a more diverse range. The integration of data products with advanced technologies enhances the precision and dependability of research findings. (7) In the context of data product application fields, Scholars in the CNKI database primarily concentrate on the application of remote sensing technology and data services, whereas scholars in the WOS database primarily focus on data processing within the realm of astronomy. Both fields share a growing focus on assessing the quality and performance of data products, as well as their utilization in disciplines like earth science, environmental science, and climate research. (8) The influence of big data and digital economy on the research of data products has become increasingly prominent, which brings new opportunities and challenges for its future research.

Based on the results and conclusions above, the following research recommendations regarding data products can be derived. (1) Enhance interdisciplinary collaboration and cross-field cooperation. Data products span across various disciplines, including computer science, environmental science, and geography. To foster innovation and enhance practical application value, it is crucial to strengthen cross-disciplinary collaboration, particularly the integration of data science with real-world applications. (2) Pay more attention to the quality and performance evaluation of data products. Both domestic and international have gradually begun to concern the research in this area. In the future, it is necessary to further strengthen research on establishing a perfect data product quality evaluation system and ensuring the accuracy and reliability of data products, so as to improve the application effect of data products in various fields. (3) Strengthen the legal and policy research of data products. As an emerging transaction object, the legal and policy environment of data products has an important impact on its development. Therefore, it is necessary to strengthen the legal and policy research of data products, clarify the property rights, transaction rules and protection measures of data products, and provide strong legal guarantee for the healthy development of data products. (4) Promote the application of data products in the field of social management and business. From the research trend perspective, data products hold significant potential for applications in social management and business sectors. The utilization of data products in these areas should be encouraged, including their use in social issue analysis, policy assessment, and business decision-making, to maximize their value. (5) Encourage cross-team cooperation and international exchanges. The research on data products still needs to strengthen cross-team, cross-institution, and industry-university-research collaborations and applications to promote rapid development and application of data products. Meanwhile, international cooperation and exchanges should also be strengthened to jointly advance research and application of data products, share experiences and technologies, and promote global development of data products.

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers for their valuable comments.

REFERENCES

[1] (2022). *The Cyberspace Administration of China, Digital China Development Report 2022*. Accessed: Feb. 25, 2024. [Online]. Available: http://www.cac.gov.cn/2023-05/22/c_1686402318492248.htm?eqid=e964285800089bd400000004646d59f6

[2] Q. Huang, Z. Zhao, and Z. Liu, "Comprehensive management system and technical framework of data quality in the data circulation transaction scenario," *Data Anal. Knowl. Discovery*, vol. 6, no. 1, pp. 22–34, Jan. 2022.

[3] R. Zhao and P. Sun, "Right confirmation, transaction price and capitalization process: Theory cogitation of big-data transformation into production factors," *J. Econ. Theory Bus. Manag.*, vol. 41, no. 1, pp. 16–26, Jan. 2021.
[4] M. Song and Z. Qin, "Reviews of foreign studies on data quality management," *J. Intell.*, vol. 26, no. 2, pp. 7–9, Feb. 2007.
[5] H. Yu and M. Zhang, "Data pricing strategy based on data quality," *Comput. Ind. Eng.*, vol. 112, pp. 1–10, Oct. 2017.
[6] J. Pei, "A survey on data pricing: From economics to data science," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 10, pp. 4586–4608, Oct. 2022.
[7] Q. Xiong and K. Tang, "Research progress on the right delimitation, exchange and pricing of data," *J. Econ.*, vol. 62, no. 2, pp. 143–158, Feb. 2021.
[8] Y. Ye, Y. Zhang, and Y. Zhu, "Exploring the form of big data products and the supporting systems," *J. Big Data*, vol. 9, no. 1, pp. 1–14, Apr. 2022.
[9] F. Li, Y. Ye, and G. Zhang, "Data value chain components and their function mechanism: A case study of digital technology promoting agricultural industry," *Sci. Technol. Manag. Res.*, vol. 42, no. 11, pp. 108–115, Jun. 2022.
[10] L. Ji, "The judicial dilemma of enterprise data protection and the dimension of breaking the situation: The road to typed right confirmation," *Leg. Forum*, vol. 37, no. 3, pp. 109–121, May 2022.
[11] R. Ouyang and W. Gong, "The pricing mechanism of data elements based on data element value and the contribution of market assessment," *Reform*, vol. 35, no. 3, pp. 39–54, Mar. 2022.
[12] R. Ouyang and Q. Du, "Research progress on the pricing mechanisms of data," *Econ. Perspect.*, vol. 63, no. 2, pp. 124–141, Feb. 2022.
[13] X. Li, "The definition and legal protection of data products," *Leg. Forum*, vol. 37, no. 3, pp. 122–131, May 2022.
[14] F. Gao and G. Ran, "A theory for data factor market formation—An mechanism framework for data factor governance," *Shanghai Econ. Res.*, vol. 41, no. 9, pp. 70–86, Sep. 2022.
[15] Y. Gao, "Theoretical justification of derivative data as the object of new intellectual property," *Soc. Sci.*, vol. 44, no. 2, pp. 106–115, Feb. 2022.
[16] Q. Huang, J. Wang, D. Chen, and X. Mo, "Research on the data price generation mechanism under the super-large scale data factor market system," *E-Gov.*, vol. 19, no. 2, pp. 21–30, Feb. 2022.
[17] Y. Wang, "An outline of the relationship between enterprise data rights and personal information protection," *J. Comparative Law*, vol. 36, no. 4, pp. 33–44, Jul. 2022.
[18] F. Gao, "Rights allocation of data holders: Legal implementation of structural separation of data property rights," *J. Comparative Law*, vol. 37, no. 3, pp. 26–40, May 2023.
[19] S. Gao, "Definition of digital labor based on Marx's view of labor," *Econ. Rev. J.*, vol. 39, no. 8, pp. 12–19, Aug. 2023.
[20] W. Liu, "On the right allocation of data products," *Peking Univ. Law J.*, vol. 35, no. 6, pp. 1581–1599, Nov. 2023.
[21] Z. Liu, "Identification of legal interests of data products and protection path of criminal law," *J. Gansu Univ. Politi. Sci. Law*, vol. 38, no. 6, pp. 74–86, Nov. 2023.
[22] J. Li and N. Gao, "Compound approach to the protection of the rights and interests of enterprise public data: Possible choice on the basis of the evolution of data type," *J. Beijing Admini. Inst.*, vol. 25, no. 2, pp. 96–108, Mar. 2023.
[23] S. Huang, "Basic theoretical issues on the 'digital economy,"' *Econ. Perspect.*, vol. 64, no. 3, pp. 3–20, Mar. 2023.
[24] Y. Sun, "Basic categories in the object of data rights and interests," *Orient. Law*, vol. 17, no. 1, pp. 24–37, Jan. 2024.
[25] (2020). *The Central People's Government of the People's Republic of China, Opinions on Building a More Complete System and Mechanism for Market-Based Allocation of Factors*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2020-04/09/content_5500622.htm
[26] (2020). *The Central People's Government of the People's Republic of China, Opinions on Accelerating the Improvement of the Socialist Market Economic System in the New Era*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2020-05/18/content_5512696.htm
[27] (2020). *The Central People's Government of the People's Republic of China, Opinions on Leading the Accelerated Development of New Consumption With New Forms and Models*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/2020-09/21/content_5545394.htm
[28] (2020). *The Central People's Government of the People's Republic of China, Instruction Opinions on Accelerating the Construction of a National Integrated Big Data Center Collaborative Innovation System*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/2020-12/28/content_5574288.htm
[29] (2021). *The Central People's Government of the People's Republic of China, Action Plan for Establishing a High-standard Market System*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2021-01/31/content_5583936.htm
[30] (2035). *The Central People's Government of the People's Republic of China, The 14th Five-Year Plan and 2035 Long-Range Goal Outline for National Economic and Social Development of the People's Republic of China*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/xinwen/2021-03/13/content_5592681.htm?eqid=945f38050007c2e300000005648fbfd0
[31] The Central People's Government of the People's Republic of China. (2021). *The 14th Five-Year*. Big Data Ind. Develop. Plan. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/2021-11/30/5655089/files/d1db3abb2dff4c859ee49850b63b07e2.pdf
[32] (2021). *The Central People's Government of the People's Republic of China, Notice on Issuing the Overall Plan for Pilot Comprehensive Reform of Factor Marketization Allocation*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/gongbao/content/2022/content_5669421.htm
[33] The Central People's Government of the People's Republic of China, Notice on Issuing. (2021). *The 14th Five-Year*. Digit. Economy Develop. Plan. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/2022-01/12/content_5667817.htm?eqid=c363a78200050cef0000000264892684
[34] (2022). *The Central People's Government of the People's Republic of China, Opinions on Speeding Up the Establishing of a National Unified Large Market*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2022-04/10/content_5684385.htm?eqid=a376b4850002f7e7000000036497a294
[35] (2022). *The Central People's Government of the People's Republic of China, Opinions on Establishing Data Basic System to Better Play the Role of Data Elements*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2022-12/19/content_5732695.htm
[36] (2023). *The Central People's Government of the People's Republic of China, Overall Layout Planning for the Construction of Digital China*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/2023-02/27/content_5743484.htm?eqid=b3ef2b740001a1ed000000066458bbec
[37] (2024). *The Cyberspace Administration of China, Notice on Issuing the 'Data Elements?' Three-year Action Plan (2024–2026)*. Accessed: Feb. 25, 2024. [Online]. Available: http://www.cac.gov.cn/2024-01/05/c_1706119078060945.htm
[38] (2023). *The Central People's Government of the People's Republic of China, Notice on Issuing the Implementation Plan for Digital Economy to Promote Common Prosperity*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/202401/content_6924631.htm
[39] (2023). *The Central People's Government of the People's Republic of China, Notice on Issuing the Instruction Opinions of Strengthening the Management of Data Assets*. Accessed: Feb. 25, 2024. [Online]. Available: https://www.gov.cn/zhengce/zhengceku/202401/content_6925470.htm
[40] J. Duan, "Health insurance data product innovation," *Chn. Finance*, vol. 71, no. 6, pp. 71–72, Mar. 2022.
[41] L. Pang, "Rethinking and reconstructing the national ownership of data resources," *J. Beijing Admini. Inst.*, vol. 24, no. 5, pp. 97–108, Sep. 2022.
[42] W. Xiao, "An analysis on legal issues of government data authorization and operation," *J. Beijing Admini. Inst.*, vol. 25, no. 1, pp. 91–101, Jan. 2023.
[43] S. Liao, J. Sun, Z. Li, L. Ma, and M. Peng, "Development, release and sharing of geoscience data products," *Adv. Earth Sci.*, vol. 20, no. 2, pp. 166–172, Feb. 2005.
[44] J. Ma, J. Li, and Z. Gao, "The introduction of quantity development of NASA earth observation science data product and its services," *Adv. Earth Sci.*, vol. 22, no. 5, pp. 521–526, May 2022.
[45] Y. An and L. Huang, "A new field of library reference: Data service," *Inform. Document Service*, vol. 29, no. 3, pp. 94–96, Mar. 2008.
[46] C. Wu, C. Lee, S. Dilshan, D. Birch, O. Tsinalis, Y. Li, S. Yan, Y. Ma, and Y. Guo, "WikiSensing: From big data to data product," *J. Fronti. Comput. Sci. Techno.*, vol. 9, no. 10, pp. 1195–1208, Oct. 2015.

[47] B. Wu and M. Zhang, "Remote sensing: Observations to data products," *Acta Geogra. Sini.*, vol. 72, no. 11, pp. 2093–2111, Nov. 2017.
[48] D. Tian, M. Gao, and C. Han, "Development implications of ESG rating and data products," *Chn. Finance*, vol. 72, no. 24, pp. 35–37, Dec. 2021.
[49] W. Liu, X. Lei, and W. Feng, "Imagine the archival data science in the era of big data," *Archiv. Manage.*, vol. 39, no. 6, pp. 55–56, Nov. 2021.
[50] H. Guo and W. N. Zhu, "A review on the spatial disaggregation of socioeconomic statistical data," *Acta Geographica Sinica*, vol. 77, no. 10, pp. 2650–2667, Oct. 2022.
[51] Y. Wan and L. Gu, "Data supply chain: From literature, intelligence to knowledge," *Inf. Stud., Theory Appl.*, vol. 46, no. 8, pp. 59–67, Aug. 2023.
[52] C. Gao, "Exploration of building industrial big data products and building new business model of sharing service in sinopec sharing center," *Financ. Accoun.*, vol. 45, no. 15, pp. 49–51, Aug. 2023.
[53] L. Xiong, M. Liu, and Z. Xu, "Research on data products pricing mechanismin China-analysis based on customer perceived value theory," *Price, Theory Pract.*, vol. 38, no. 4, pp. 147–150, Apr. 2018.
[54] S. Luo and L. Xing, "Neutrosophic game pricing methods with risk aversion for pricing of data products," *Expert Syst.*, vol. 38, no. 5, Aug. 2021, Art. no. e12697.
[55] S. Luo, W. Pedrycz, and L. Xing, "Pricing of satellite image data products: Neutrosophic fuzzy pricing approaches under different game scenarios," *Appl. Soft Comput.*, vol. 102, Apr. 2021, Art. no. 107106.
[56] H. Yu and J. Huang, "Data product pricing strategy based on closed-loop data supply chain," *J. Indus. Eng. Eng. Manag.*, vol. 37, no. 1, pp. 136–146, Jan. 2023.
[57] X. Guo, Q. Li, and H. Wang, "Analysis of sampling strategy for data product trading under perceived value uncertainty," *Fronti. Sci. Techno. Eng. Manage.*, vol. 42, no. 2, pp. 25–33, Mar. 2023.
[58] J. Wang, Y. Dou, L. Huang, and G. Li, "Data product pricing: A review of research progress and comparison of pricing methods," *Price Theory Practi.*, vol. 43, no. 4, pp. 22–27, Apr. 2023.
[59] W. Pan, L. Xiao, R. Zhan, and Y. Ye, "Research on valuation and pricing models for public and enterprise data-exploration of Guizhou practice based on data product trading price calculator," *Price Theory Pract.*, vol. 43, no. 8, pp. 44–50, Aug. 2023.
[60] X. Guo, Q. Li, H. Wang, and J. Du, "Auction pricing mechanism of data transactions under demand information asymmetry," *Operat. Res. Manage. Sci.*, vol. 32, no. 11, pp. 170–175, Nov. 2023.
[61] Z. Wang, Q. Huang, and M. Ren, "Expected income distribution model of data trading in federated computing," *J. Mod. Inform.*, vol. 44, pp. 1–16, Feb. 2024.
[62] X. Zhang, Y. Dou, C. Zhang, and L. Huang, "Research on the profit-sharing mechanism for federated learning," *Fronti. Sci. Techno. Eng. Manage.*, vol. 42, no. 2, pp. 8–15, Mar. 2023.
[63] X. Xu, "Challenges to government statistics brought by the development of digital economy," *J. Statis. Inf.*, vol. 38, no. 10, pp. 3–8, Oct. 2023.
[64] S. Li and J. Yao, "A characteristics and assessment analysis of dem products," *Prog. Geogra.*, vol. 24, no. 6, pp. 99–108, Dec. 2005.
[65] X. Zhang, X. Li, H. Zhang, and L. Wang, "Research on GIS data product quality evaluation based on fuzzy gravity center comprehensive evaluation," *Sci. Srvyin. Mapin.*, vol. 32, no. 6, pp. 49–51, Dec. 2007.
[66] J. Wang, J. Zhang, and J. Wang, "Quality assessment of spaceborne microwave radiometer aquarius data product based on Argo Buoy data," *Haiyang Xuebao*, vol. 37, no. 3, pp. 46–53, Mar. 2015.
[67] F. Chen, "Research on selection of data products based on hesitant linguistic judgment matrixs," *Comput. Eng. Appl.*, vol. 53, no. 15, pp. 95–100, Aug. 2017.
[68] Y. Dai, "Selection of data product service provider based on hesitant fuzzy linguistic decision-making model," *Comput. Eng. Appl.*, vol. 54, no. 12, pp. 133–137, Jun. 2018.
[69] Y. Liang, "Selection of data products based on probabilistic hesitant fuzzy information aggregation algorithm," *Comput. Eng. Appl.*, vol. 55, no. 3, pp. 219–224, Feb. 2019.
[70] X. Lin, "Data product selection based on linguistic information consistency adjustment algorithm," *Comput. Eng. Appl.*, vol. 55, no. 21, pp. 129–134, Nov. 2019.
[71] E. Shi, "The selection of data product service provider based on the single-valued neutrosophic model," *Control Eng. China*, vol. 27, no. 2, pp. 391–395, Feb. 2020.
[72] S. Luo, W. Pedrycz, and L. Xing, "Selection of data products: A hybrid AFSA-MABAC approach," *Int. J. Mach. Learn. Cybern.*, vol. 13, no. 4, pp. 1079–1097, Apr. 2022.
[73] X. Li and H. Li, "A visual analysis of research on information security risk by using CiteSpace," *IEEE Access*, vol. 6, pp. 63243–63257, 2018.
[74] Z. Zeng and T. Hengsadeekul, "Environmental issues and social responsibility: A scientomeric analysis using citespace," *Entrepreneurship Sustainability Issues*, vol. 8, no. 2, pp. 1419–1436, Dec. 2020.

YANNI LIANG was born in China, in 1990. She received the B.S. and Ph.D. degrees from North China Electric Power University, Beijing, in 2013 and 2018, respectively.

She is currently a Postdoctoral Researcher with the School of Economics and Management, Tongji University, and a special Teacher with the School of Economics and Management, Beibu Gulf University. Her research interests include digital economy and supply chain management.

JIANXIN YOU was born in China, in 1961. He received the M.S. and Ph.D. degrees from Tongji University, in 1992 and 1999, respectively.

He is currently an Academician with the International Academy for Quality and a Professor with the School of Economics and Management, Tongji University. His research interests include digital economy and quality management.

RAN WANG was born in China, in 1990. He received the B.S. and M.S. degrees from North China Electric Power University, Beijing, in 2013 and 2016, respectively.

He is currently with SkyCloud Technology Company Ltd. His research interests include cloud computing technology and digital economy.

SHUO HAN was born in China, in 1997. He received the B.S. degree from Inner Mongolia University of Science and Technology, in 2016. He is currently pursuing the master's degree with the School of Economics and Management, Beibu Gulf University.

His research interests include cloud digital trade and regional economies.