Analysis of Author Collaboration and Research Content on Open Sharing of Chinese Scientific Data: A Postprint

Authors: Sheng Xiaoping, Sun Qianqian

Date: 2023-04-01T16:02:59+00:00

Abstract

[Purpose/Significance] To reveal author collaboration relationships and collaborative research content on the topic of scientific data open sharing in China, thereby providing references for broader scientific data open sharing. [Method/Process] Utilizing journal and conference paper data on scientific data open sharing from CNKI between 2004 and 2021, this study statistically analyzes the basic situation of author collaboration, constructs an author collaboration network, and employs social network analysis methods to examine the overall author collaboration network, collaboration groups, and collaborative research content. [Results/Conclusions] Journal and conference papers on scientific data open sharing in China are primarily based on author collaborative research. The overall network density of author collaboration is relatively low; both the scale of collaboration groups and the frequency of collaboration influence authors' centrality in the network. The structure of author collaboration groups exhibits five patterns: dual-core, network-frame, bridge-connection, star, and streamlined types. Author collaborative research content covers nine aspects: big data, scientific data, open sharing, data publishing, data management, data policy, data security, government data, and data governance.

Full Text

Author Cooperative Relationships and Cooperative Research Content on the Topic of Domestic Scientific Data Open Sharing

Sheng Xiaoping, Sun Qianqian

School of Library, Information and Archives, Shanghai University, Shanghai 200444, China

Abstract:
[Purpose/Significance] This study aims to reveal author cooperative relationships and cooperative research content on the topic of domestic scientific data open sharing, in order to provide references for broader scientific data open sharing. [Method/Process] Based on journal and conference paper data from CNKI on the topic of scientific data open sharing from 2004 to 2021, we analyzed the basic situation of author cooperation and constructed author cooperation networks. We then analyzed the overall author cooperation network, cooperative groups, and cooperative research content using social network analysis methods. [Result/Conclusion] Journal and conference papers on domestic scientific data open sharing are primarily based on author cooperative research. The overall network density of author cooperation is low, and both the size of cooperative groups and frequency of cooperation affect author centrality in the network. Author cooperative group structures present five patterns: double-core, grid, bridge-connection, star, and streamline. Author cooperative research content covers nine aspects: big data, scientific data, open sharing, data publishing, data management, data policy, data security, government data, and data governance.

Keywords: scientific data; open sharing; author cooperation; cooperative relationship; social network
Classification Number: G203
DOI: 10.13266/j.issn.0252-3116.2021.23.002

With the development of data-intensive scientific paradigms, open sharing of scientific data has gradually become a consensus in the academic community. Meanwhile, to reduce scientific costs and accelerate scientific output, researchers increasingly need to conduct collaborative research. Current scholarship has extensively explored aspects such as scientific data open sharing policies [1], responsibilities and participation motivations of various stakeholders [2], service innovation and evaluation [3], platforms and applications [4], operation and guarantee mechanisms [5], and countermeasures and suggestions [6]. However, few studies have examined author cooperative relationships and cooperative research content on the topic of domestic scientific data open sharing. This paper attempts a preliminary analysis of this issue to further promote the development of domestic scientific data open sharing practice.

We retrieved data from the CNKI database, limiting document types to journals and conferences. Since dissertations are completed by single authors and do not involve inter-author cooperation, they were excluded from this study. Using the search terms "scientific data" or "research data" combined with "open sharing" (search date: September 26, 2021), we obtained 1,016 relevant papers, including 1,000 journal papers and 16 conference papers. After removing anonymous authors (those without author identification and newsletter-type articles) and duplicate papers, and replacing author names listed as "Research Group" with actual group members, we obtained 884 papers (868 journal papers and 16 conference papers) as our sample for further analysis.

Literature retrieval revealed that as early as 2004, the topic of scientific data open sharing had already attracted attention in China. However, from 2004 to 2012, the number of publications on this topic remained in single digits. After 2012, the number of published papers, co-authored papers, and contributing authors all showed a growth trend. This analysis focuses on collaboration as reflected in co-authorship. Among the 884 papers, 492 were completed through multi-author collaboration. The total number of authors reached 1,518, including 321 independent authors (who did not participate in collaborative publishing) and 1,197 co-authors. In terms of single-paper collaboration, 2-3 author collaborations were most common, while papers with seven or more co-authors were rare, with one paper having as many as 23 co-authors (see Table 1). In short, the collaboration degree for papers on this topic is 1.72 (collaboration degree = total number of authors / total number of papers × 100%), and the collaboration rate is 55.67% (collaboration rate = number of collaborative papers / total number of papers × 100%). Co-authors account for 78.85% of all authors on this topic, indicating that most domestic journal and conference papers on scientific data open sharing are collaborative research.

1 Data Sources and Research Methods

This study primarily employs social network analysis methods and tools such as VOSviewer and UCINET to analyze author cooperation on the topic of domestic scientific data open sharing from three perspectives: overall author cooperation network, author cooperative groups, and cooperative research content.

2 Overall Author Cooperation Network Analysis

Treating authors as nodes in the cooperation network and their joint authorship relationships as edges (without weighting contributions within a single paper), we constructed an undirected symmetric overall network of author cooperation on domestic scientific data open sharing (see Figure 1 [Figure 1: see original paper]), where a, b, and c represent the top three groups in terms of number of authors with cooperative relationships. Group a consists of 102 people, group b consists of 29 people, and group c consists of 28 people. The overall author cooperation network is a primary way to reflect cooperative relationships, and its density and centrality analyses are as follows.

2.1 Density Analysis

Network density reflects the sparsity of relationships between nodes, with values ranging between 0 and 1. Values closer to 1 indicate tighter relationships [7]. From 2004 to 2021, the density of the overall author cooperation network on domestic scientific data open sharing is only 0.004, with a standard deviation of 0.076 for relationships in the network. The low network density indicates a dispersed overall network with weak connectivity, meaning relationships between nodes are too sparse and suggesting that communication and interaction among authors face certain obstacles.

2.2 Centrality Analysis

Centrality analysis includes measurements of degree centrality, betweenness centrality, and closeness centrality of nodes in the network. When calculating degree centrality, since one paper was co-authored by 23 people, and among them Hao Jinxin, Xue Yanjie, and Cui Chenzhou also co-authored another paper with three other authors, the degree centrality of these 23 authors is above 22. Using Price's Law $K = 0.749\sqrt{Nmax}$ to screen core authors with high collaborative output, where Nmax represents the maximum number of collaborative papers. In our sample data, the maximum number of collaborative papers is 17, yielding $K = 3.09$. Thus, core authors in scientific data open sharing are those with 4 or more collaborative papers, totaling 26 authors (see Table 2). However, the top three authors by collaborative output—Huang Ruhua, Gu Liping, and Sheng Xiaoping—have degree centrality values of 15, 19, and 13 respectively, which are lower than the aforementioned 23 authors. This is a special case resulting from the relatively small scale of the author cooperation overall network and insufficient sample size, but degree centrality shows a positive correlation with cooperative group size.

A higher betweenness centrality value indicates that a node lies on more network paths between other nodes [8]. In Table 2 , Shi Lei has the highest betweenness centrality value, meaning she plays the most significant bridging role in the entire network and has the strongest ability to control resources. Among the top 26 authors by betweenness centrality, Huang Ruhua and Huang Yuting belong to the 29-person group, Yang Fan belongs to the 28-person group, and the remaining 23 authors all belong to the 102-person group (see Figure 1 [Figure 1: see original paper]). However, among the top 26 authors by collaborative output, Sheng Xiaoping, Ma Haiqun, Xing Wenming, Wen Fangfang, Xu Ting, Qian Qing, Yu Guangjun, Li Yang, Wu Sizhu, Zhang Xiaoyue, Yang Xianmin, and Zhu Yangyong do not appear among the top 26 authors by betweenness centrality. This demonstrates that while there is a positive relationship between the number of co-authored papers and author betweenness centrality, it is not a simple linear relationship. Author betweenness centrality also has a positive relationship with the size of the cooperative group to which the author belongs.

A smaller closeness centrality value indicates shorter distances between a node and other nodes, making information acquisition and transmission easier and less dependent on others. The top 26 closeness centrality values are all relatively large (see Table 2), and all top 26 authors by closeness centrality belong to the largest 102-person group (see Figure 1 [Figure 1: see original paper]), including Shi Lei, Gao Mengxu, Wang Ruidan, Wang Chao, Hu Lianglin, Xu Bo, Gu Liping, Li Jianhui, Zhang Lili, Wen Liangming, Wang Jian, Zhou Yuanchun, Li Chengzan among the top 26 by collaborative output. However, the remaining 13 authors from the top 26 by output do not appear among the top 26 by closeness centrality. This indicates that information transmission in the overall author cooperation network on domestic scientific data open sharing is not yet smooth enough, though relatively good information access pathways have formed within cooperative groups. Therefore, in the author cooperation network for domestic scientific data open sharing, both the size of the cooperative group to which an author belongs and the frequency of cooperation affect their betweenness centrality and closeness centrality.

3 Author Cooperative Group Analysis

Author cooperative groups are another primary parameter reflecting author cooperation relationships. The overall author cooperation network structure shown in Figure 1 [Figure 1: see original paper] is dispersed, with most groups being independent of each other and cross-group cooperation being rare. Therefore, it is necessary to further analyze the quantity and structure of author cooperative groups.

3.1 Quantity Analysis of Author Cooperative Groups

The 1,197 participating authors formed 314 cooperative groups. The most common are small groups of 2 authors, totaling 150 groups, followed by small groups of 3 authors, totaling 84 groups. There are 9 unique cooperative groups (see Table 3). The largest author cooperative group has 102 people, formed by 6 small groups through "cooperative relationships" (see Figure 2 [Figure 2: see original paper] left). The second-largest author cooperative group has 29 people, formed by 2 small groups through "cooperative relationships" (see Figure 2 [Figure 2: see original paper] right). Thus, the domestic scientific data open sharing research topic has initially formed a certain number of cooperative groups of varying sizes.

3.2 Structure Analysis of Author Cooperative Groups

This study summarizes the topological structure types in the author cooperation network on scientific data open sharing into five patterns: double-core, grid, bridge-connection, star, and streamline [9]. The double-core structure refers to two-person cooperation groups (see Figure 3 Figure 3: see original paper). The grid structure refers to author cooperation groups where any two nodes are connected (see Figure 3 Figure 3: see original paper). The bridge-connection structure refers to author cooperation groups that require bridge nodes to connect multiple nodes or cooperation groups, presenting a mixed network structure state (see Figure 3 Figure 3: see original paper). The star structure refers to author cooperation group patterns where one core node connects dispersed nodes (see Figure 3 Figure 3: see original paper). The streamline structure refers to author cooperation group patterns where nodes form cooperation relationships in a linear rather than cross manner (see Figure 3 Figure 3: see original paper). Among the five patterns, the double-core structure is most common (accounting for 47.8%), followed by the grid structure (39.8%). The streamline structure only appears in small cooperation groups of 3-5 people, while cooperation groups with more than 10 people all present bridge-connection structures (see Table 3).

4 Author Cooperative Research Content Analysis

We selected 492 cooperative papers for author cooperative research content analysis, extracted 1,158 keywords, merged synonymous keywords such as "scientific data" and "research data," "government data opening" and "open government data," etc. After constructing a keyword matrix and importing it into VOSviewer, we selected high-frequency keywords appearing 7 or more times for clustering, generating a high-frequency keyword co-occurrence map (see Figure 4 [Figure 4: see original paper]). This revealed nine sub-themes: big data, scientific data, open sharing, data publishing, data management, data policy, data security, government data, and data governance. The cooperative research content is as follows:

4.1 Big Data

This theme focuses on the big data context of scientific data open sharing, big data fields, big data technologies, and big data requirements for open sharing. In the era of big data, both data processing and application methods have undergone tremendous changes. Scientific research has entered a new paradigm characterized by data-intensive scientific discovery, continuously increasing demands for scientific data preservation, management, and sharing [10], while generating corresponding open sharing needs and diverse open sharing models in strategic, public, and numerous other fields. People pay special attention to scientific data resource status and sharing services in health and medical big data, agricultural big data, earth observation big data, credit big data, and government big data. However, since traditional data processing software cannot process various types of big data within specified time limits, big data technologies are needed, such as big data image recognition technology, video recording technology, and information collection technology, combined with new-generation information and network technologies like the Internet of Things, cloud storage, and artificial intelligence to jointly facilitate the sharing of scientific and technological resources, enabling data integration, mining, display, and management [11]. Therefore, the big data era requires strengthening quality management of scientific data to ensure the standardization, accuracy, authenticity, and timeliness of openly shared scientific data; establishing institutional alliance big databases to strengthen data information sharing and interactive connectivity among similar institutions [12]; and promoting cross-integration and comprehensive application of scientific data from different types and disciplines.

4.2 Scientific Data

This theme focuses on the connotation and characteristics, current status and problems, management policies, influencing factors, and coping strategies of scientific data. Scientific data refers to various basic and observational data materials and related information generated in scientific and technological activities. It is the most important information carrier of scientific achievements, with characteristics of objectivity, diversity, basicness, resourcefulness, transmissibility, shareability, and value-addedness [13]. Domestic universities and research institutions have successively established data repositories, publishers have launched data journals, and commercial companies have introduced various data services, with the number of scientific datasets growing exponentially [14]. However, compared with foreign countries, domestic scientific data sharing platform construction started relatively late, laws and regulations are missing, and the institutional guarantee system remains to be perfected; integration is limited, and data sources are relatively single. Domestic and foreign scientific data open sharing is influenced by eight aspects: scientific personnel, policy, data, technology, organization, platform, law, and funding [15]. To further promote scientific data open sharing, we should build an open sharing system by improving laws, regulations, and standard systems, perfecting management planning and supervision models, using new-generation information technology to achieve technical guarantees, strengthening internal and external cooperation and collaborative construction, and enhancing talent team building [16].

4.3 Open Sharing

This theme concentrates on principles, sharing models, operation models and implementation approaches, existing problems, key mechanisms, and strategies of scientific data open sharing. FAIR (Findable, Accessible, Interoperable, Reusable) has become a widely accepted and recognized principle for scientific data open sharing in the scientific community [17]. Current scientific data sharing models at home and abroad can be summarized into five types: policy-led, organization-cooperation, topic-driven, data-publishing, and data-market [18]. The operation model of scientific data open sharing usually includes government-led public welfare operation models and market-led commercial operation models, with three main implementation approaches: data repositories, data journals, and data description documents in academic journals [19]. However, China's scientific data open sharing is still in a basic, theoretical exploration stage, lacking strong policy and regulation support, with relatively little business communication and project cooperation with international organizations; uneven distribution of scientific and technological data resources and unbalanced sharing progress; complicated construction of scientific and technological resource sharing platforms lacking unified standards; and strong traditional conservative ideas restricting sharing enthusiasm [20]. Therefore, we need to continuously strengthen data open sharing drivers through effective incentive mechanisms, improve data open sharing models through long-term evaluation mechanisms, and expand open data dissemination channels through multi-level propagation mechanisms to achieve a virtuous cycle of the complete chain from data opening to dissemination to reuse. Additionally, China urgently needs to adopt effective strategies to promote scientific data open sharing, including building a scientific data open sharing system, strengthening scientific data preservation, accumulation, analysis, and mining, reasonably defining and effectively protecting scientific data rights, gradually promoting graded and classified opening of scientific data, enhancing international cooperation practices in scientific data open sharing [21], and implementing open sharing by developing standardized data products [22].

4.4 Data Publishing

This theme focuses on the role and requirements of scientific data publishing, publishing models and processes, journal data publishing policies, and problems and countermeasures facing data publishing. Data publishing can not only promote data open sharing and facilitate the development of academic publishing and scientific research but also utilize data open sharing to solve data intellectual property issues, ensure scientific data quality, and improve the value of data reuse [23]. Scientific data publishing needs to meet the requirements of standardized citation, unique identification, and convenient utilization of scientific data [15]. Data publishing models include three types: independent data publishing (i.e., storage and release in data repositories), data release as supplementary materials to papers, and release in the form of data papers [24]. The general data publishing process includes five stages: data submission, data storage, data review, data citation, and data evaluation [25]. Currently, data publishing policies of major international publishing institutions involve four aspects: data submission, data review, data opening and storage, and data protection, actively encouraging authors to submit and share data; while domestic journal data policies' requirements and explanations for data publishing are not yet systematic [26]. The main problems facing data publishing are that compared with traditional publishing, data publishing still needs improvement in establishing academic status and industry recognition, with immature operation models [27], insufficient data storage infrastructure construction, imperfect data review systems, serious data intellectual property infringement, and poor data citation effects [28]. To address these challenges, China's scientific data publishing should accelerate the establishment of a batch of open, standardized, strictly quality-controlled data repositories supporting FAIR principles; explore new publishing models and establish rights protection mechanisms in data publishing; standardize data citation; perfect data review systems to ensure scientific data quality; attach importance to data intellectual property; and promote digital asset management [29].

4.5 Data Management

This theme focuses on objects and principles of data management, scientific data management planning and methods, scientific data management institutions and services, practical status, and development suggestions. The objects of data management mainly include big data, government data, scientific data, and enterprise data [30], with principles mainly including data value quantification, data quality control, metadata management, planning management, team collaboration, dynamic management, lifecycle management, risk management, technical relevance, and leadership support [31]. Scientific data management planning describes what data will be produced and how data will be managed throughout its lifecycle and made accessible, including project-expected data content, type, scale, quality, submission time, and the name of the final scientific data management institution, helping to advance scientific data management in a planned manner. The "Scientific Data Management Measures" provides macro guidance for China's scientific data management, with its content, implementation status, influencing factors, and promotion strategies receiving attention [32]. Scientific data management units at all levels need to clarify responsibilities and collaboratively participate in scientific data management practice according to the "Scientific Data Management Measures"; library and information institutions should vigorously promote scientific data management services [33]; and data management institutions should widely apply new technologies to actively promote the maturity and development of open data management. Domestic scientific data management practice is becoming increasingly mature, but scientific data management mostly concentrates on natural and engineering sciences, with much room for improvement in social science data management practice [34]. The following measures can be taken to promote China's scientific data management: formulate highly operational implementation plans and implement data sharing methods; strengthen scientific data sharing supervision, intellectual property protection, and research and regulation of cross-border data management; enhance national scientific data center construction and continuous scientific data integration; raise data management awareness; and cultivate relevant talents [27].

4.6 Data Policy

This theme focuses on investigating and analyzing data policies of domestic and foreign organizations or institutions. Many international organizations have formulated scientific data open sharing policies that emphasize the value of scientific data open sharing, intellectual property and its protection, interoperability and high quality, and evaluation and stakeholder responsibility [35]. International publishing institutions, scientific education institutions, and scientific funding institutions have also formulated relevant scientific data policies involving deposit policies, storage specifications, supervision mechanisms and quality assurance specifications, openness degree, dissemination specifications, and utilization methods. For example, SpringerNature has formulated four different levels of data licensing policies: encouraging, strongly recommending, partially mandatory, and fully mandatory [36]. China has initially built a policy system for scientific data management and sharing from central to local levels, but the status of regional open data policy construction is uneven, and local government open data policies are mainly issued in the form of opinions, notices, and plans, with policy binding force and enforceability needing improvement and policy implementation requiring continued accumulation [37]. Certain data policy content such as how to divide data rights and interests and how to implement network security deserve further research and discussion [38].

4.7 Data Security

This theme emphasizes laws, regulations, policies, technologies, and governance of data security. Data security involves relevant laws and standards, as well as technologies such as data encryption, software and hardware data protection mechanisms, backup, data masking, and data clearing [39]. Before the promulgation of the "Data Security Law of the People's Republic of China," China's relevant laws for ensuring data security mainly included the "Cybersecurity Law of the People's Republic of China" and the "National Security Law of the People's Republic of China" [40]. Currently, China's data security policies focus on network security protection, personal privacy protection, data cross-border flow, and security protection of open data platforms [41], while also paying attention to data security governance. In the process of implementing data security technical governance, we should make full use of blockchain technology [42], privacy-enhancing technology, data authentication technology, and building national scientific data centers based on data alliances [43]. However, maintaining data security is not only a technical issue but also a management issue [44]. Therefore, we need to strengthen data security management governance, including establishing scientific data classification and grading standards and systems, implementing data protection impact assessments, and improving scientific data quality [15]. In the long run, data security is certainly important, but data open sharing is the general trend, so we need to improve data open guarantee mechanisms and find a balance between data opening and data security [45].

4.8 Government Data

This theme focuses on the value, policy and standard systems, problems, and countermeasures of government data opening. Government data opening mainly refers to government agencies opening and sharing datasets they produce or own within the scope of laws and regulations, which can be freely used by enterprises, citizens, researchers, etc., to create value for society [46]. Government data opening has political, social, economic, and technical value [47]. Since the State Council issued the "Outline for Promoting Big Data Development" in 2015, central and local governments have formulated a series of data opening policies involving data infrastructure implementation policies, data resource management policies based on the open government data lifecycle, and government governance policies [48]. Domestic scholars are also interested in foreign government data opening policies, regulations, and standards [49], as well as cooperation models and construction experience [50]. Government data opening and sharing need to pay attention to government data and metadata, data management, data security and privacy, data rights and interests, and user participation. To address these issues, we need to further standardize government data opening approaches, strengthen copyright protection of government works, improve functions of local government data opening platforms, and focus on research on China's government data opening standards, data quality control, and user utilization [51].

4.9 Data Governance

This theme focuses on the definition, scope and types, existing problems, and governance ideas and suggestions of data governance. Data governance is the process of formulating standards, rules, policies, and implementing supervision to ensure compliance with data management best practices [31]. In terms of governance scope, data governance can be divided into global data governance, national data governance, government data governance, and enterprise data governance [52]. Since China's relevant theoretical system of data governance is not yet complete, many challenges and obstacles have been encountered in the data governance process, with problems existing in institutional design, technical capabilities, human resources, management, and data [53], such as lagging key big data technology research and development, unclear levels of data governance systems, vague data sovereignty protection rights and responsibilities, difficulties in realizing data production factor value, and risks to data security and personal privacy [54]. To further implement data governance, we should improve data quality, establish data transaction circulation systems, strengthen data talent cultivation and data governance system construction, and set up data governance steering committees and data governance advisory committees [55].

Conclusion

Through analysis of the overall author cooperation network, quantity and structure of author cooperative groups, and author cooperative research content on domestic scientific data open sharing, we find that: (1) After 2012, the number of published papers, co-authored papers, and authors on domestic scientific data open sharing all showed a growth trend, with more than half of journal and conference papers being collaborative research outcomes. (2) The overall network density of relevant author cooperation on domestic scientific data open sharing is low, with strong dispersion and weak connectivity, indicating that scientific data sharing and communication interaction among these authors face certain obstacles. Both the size of the cooperative group to which authors belong and cooperation frequency affect their betweenness centrality and closeness centrality in the author cooperation overall network. (3) The domestic scientific data open sharing research topic has formed 314 small cooperative groups comprising 1,197 co-authors, including 150 two-person groups, 84 three-person groups, and other small groups, with the largest small group consisting of 102 people. These cooperative groups present five topological structure patterns in the scientific data open sharing topic author cooperation network: double-core, grid, bridge-connection, star, and streamline, among which double-core and grid structures are the main patterns. (4) Author cooperative research content on domestic scientific data open sharing mainly covers nine sub-themes: big data, scientific data, open sharing, data publishing, data management and governance, data policy, data security, government data, and data governance, forming a series of relevant theoretical viewpoints. With further development of scientific data open sharing, new types of open sharing cooperative relationships and cooperative research content may emerge in scientific research, which is worth continued attention in future research.

References

[1] Ma Haiqun, Pan Pan. Analysis of current research status of open data policies at home and abroad and research on our country's research…

[2] Sheng Xiaoping, Wu Hong. Dynamics of different stakeholders in scientific data open sharing activities…

[3] Wu Tong. Research on open sharing of scientific data in American research libraries based on data lifecycle…

[4] Zhou Honghong, Wu Shiyu. Current status of construction of scientific and technological information resource sharing platforms in China…

[5] Wei Junchao, Yan Yan, Wang Haibiao. Construction and application of scientific data open policy guarantee system…

[6] Zhang Xiaoyue, Gu Liping, Hu Lianglin. Review of solutions to the dilemma of reusing open scientific research data at home and abroad…

[7] Liu Jun. Introduction to Social Network Analysis [M]. Beijing: Social Sciences Academic Press, 2004: 100.

[8] Qiu Junping, Wang Feifei. Research on author cooperation relationships in domestic competitive intelligence field based on SNA…

[9] Liu Bei, Yuan Yi, BOUTINE. Application research of social network analysis method in paper cooperation networks…

[10] Zhao Ruixue, Zhao Hua, Zhu Liang. Progress in construction and sharing of agricultural scientific big data at home and abroad…

[11] Li Jia, Wang Hongqi, Li Yue, et al. Evolutionary game research on technology resource sharing behavior among regional innovation service platforms in the era of big data…

[12] Li Shewang, Liu Danting. Realistic dilemmas and path transformation of sports think tank construction in China from the perspective of big data…

[13] Ma Ning, Liu Zhao. Research on open sharing system of scientific data resources…

[14] Luo Pengcheng, Cui Haiyuan, Zhao Jingru. Research on current character-

istics of scientific data based on DataCite…
[15] Sheng Xiaoping, Yuan Yuan. Review of influencing factors of scientific data open sharing at home and abroad…
[16] Li Yajing, Ma Haiqun. Review of scientific data research in China based on bibliometrics…
[17] Song Jia, Wen Liangming, Li Yang. FAIR principles for scientific data sharing: background, content, and practice…
[18] Li Shanqing, Xing Xiaozhao. Research on scientific data sharing models in the context of big data…
[19] Liu Jingjing, Ma Jianhua. On three approaches to open sharing of scientific research data…
[20] Zhang Wang, Cheng Huiping. Research on strategies, mechanisms, and optimization paths of scientific data open sharing…
[21] Li Sijing, Song Lirong, Wang Jian. Scientific data publishing for open sharing: opportunities, challenges, and countermeasures…
[22] Lan Fangxin, Wu Xingyang, Xu Ping. Extraction and product development of precipitation data in ground meteorological observation monthly data files…
[23] Chu Jiewang, Wang Min. Enlightenment of American scientific data open sharing strategies to China…
[24] Li Chengzan, Zhang Lili, Hou Yanfei, et al. Open sharing of scientific big data: models and mechanisms…
[25] Yang Jing, Kang Qi, Li Zhe. Thoughts and enlightenment on promoting scientific data open sharing…
[26] Liu Ziheng, Tu Zhifang. Review of data publishing and its quality control…
[27] Qu Baoqiang, Song Lirong, Wang Jian. Development trends of scientific data publishing from the perspective of open sharing…
[28] Kong Lihua, Xi Yan, Zhang Xiaolin. Trends, mechanisms, and challenges of data publishing…
[29] Xu Jie, Tang Wenhui, Xia Xinyue. Analysis of current situation and countermeasures of data publishing oriented to practice…
[30] Huang Haiying, Xiao Lan, Wang Bingjie, et al. Research progress and hotspots of data management: strategy, roles, and applications…
[31] Sheng Xiaoping, Song Dacheng. Comparative analysis of data management and data governance and its enlightenment for formulating scientific data open sharing policies…
[32] Xing Wenming, Hong Fanglin, Li Xiaoyan. Two-dimensional perspective of scientific data management system—Interpretation of "Scientific Data Management Measures"…
[33] Qiu Chunyan, Huang Ruhua. New progress in international scientific data sharing field in recent 3 years…
[34] Zhang Lili, Wen Liangming, Shi Lei, et al. Latest progress in scientific data management and open sharing at home and abroad…
[35] You Xiaguang, Sheng Xiaoping. Comparison and characteristic analysis of scientific data open sharing policies of 8 international organizations…
[36] Liu Fenghong, Peng Lin. Data policies and practices of international publishing groups under the background of FAIR principles…

[37] Tan Biyong, Liu Rui. Research on local government open data policies in China—Taking 15 sub-provincial cities as examples…
[38] Gu Liping, Bao Qinwen. Incentive policies for open scientific research data abroad…
[39] Ding Hongfa, Meng Qiouqing, Wang Xiang, et al. Analysis of data security and privacy protection countermeasures for government data opening oriented to data lifecycle…
[40] Sheng Xiaoping, Guo Daosheng. Research on data security governance in scientific data open sharing…
[41] Huang Ruhua, Miao Miao. Security protection countermeasures for Chinese government open data…
[42] Gao Yongchang, Gao Jing, Li Deqiang. Application practice of blockchain in social security data assets and sensitive information security protection…
[43] Li Shanqing, Zheng Yanning, Xing Xiaozhao, et al. Research on security management issues of scientific data sharing…
[44] Ma Haiqun, Zhang Tao, Li Zhongjun. System dynamics of government data opening and security under COVID-19…
[45] Sheng Xiaoping, Wang Yi. Responsibilities and roles of stakeholders in scientific data open sharing…
[46] Xiao Min, Guo Qiuping, Mo Zuying. Difference analysis of government data opening development history and platform construction…
[47] Huang Ruhua, He Naidong, Li Baiyang. Construction of value system of China's open government data…
[48] Huang Ruhua, Wen Fangfang, Huang Wen. Construction of China's government data open sharing policy system…
[49] Zhao Rongying, Liang Zhisen, Duan Peipei. Metadata standards for UK government data opening and sharing—Investigation and enlightenment of Data.gov.uk…
[50] Wu Lin, Wu Shiyu. European open government data cooperation model and implementation—Cross-regional co-construction and sharing model…
[51] Ma Wucui, Liu Wenyun, Su Qingshou, et al. Analysis of current situation and development countermeasures of local government data opening in China…
[52] Yang Lin, Si Mengmeng, Zhu Yangyong. Exploration of multi-level data governance system framework construction ideas…
[53] Yang Jing, Kang Qi, Li Zhe. Analysis and enlightenment of U.S. Federal Data Strategy and 2020 Action Plan…
[54] Zheng Yueping, Gan Qixuan, Zhang Caiwei, et al. Current situation and problems of local government data governance—Empirical research based on 43 government hotline departments…
[55] Yin Guowei, Nie Fengying, Du Shaoming. Online open sharing of U.S. agriculture-related data…

Author Contributions

Sheng Xiaoping: Designed research framework, revised paper;
Sun Qianqian: Collected and processed data, wrote paper.

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