Postprint Review of Scientific Data Research in Domestic Libraries

Authors: Chai Huiming, Zhang Libin, Zhao Yajie

Date: 2023-07-26T00:00:00+00:00

Abstract

[Purpose/Significance] To review and analyze the current state of domestic library science data research, summarize its characteristics, and propose recommendations for future studies, thereby enabling library science data research to better serve scientific research development and facilitate the transformation of library services. [Method/Process] This study employs a combined research methodology of bibliometric analysis and thematic analysis, with emphasis on the latter. It examines three fundamental concepts—data-intensive research, scientific data, and scientific data curation—and summarizes the research characteristics. [Results/Conclusion] Current domestic research demonstrates three fundamental characteristics: (1) emphasis on scientific data issues; (2) gradual deepening of research; and (3) attention to foreign experiences with weak localization research. Future research directions should focus on: (1) enriching research methods and strengthening empirical research; (2) promoting interdisciplinary research; and (3) enhancing fundamental research suited to the development needs of data-intensive research.

Full Text

A Review of Research on Scientific Data in Domestic Libraries

Chai Huiming[1], **Zhang Libin**[2], **Zhao Yajie**[2] [1] Library of Tianjin Conservatory of Music, Tianjin 300171 [2] Nankai University Library, Tianjin 300071

Abstract:

[**Purpose/Significance**] This article systematically reviews and analyzes the current state of research on scientific data in domestic libraries, summarizes its key characteristics, and offers recommendations for future research to better serve scientific development and promote the transformation of library services.

[**Method/Process**] The study employs a combination of bibliometric analysis and thematic analysis, with particular emphasis on the latter, examining three core concepts: data-intensive research, scientific data, and data curation. [**Results/Conclusions**] Current domestic research exhibits three fundamental characteristics: (1) emphasis on scientific data issues; (2) gradual deepening of research; and (3) focus on foreign experiences with weak localization research. Future research should: (1) enrich research methods and strengthen empirical studies; (2) promote interdisciplinary research; and (3) enhance basic research suited to the development needs of data-intensive research.

Keywords: data-intensive research; scientific data; data curation; research data services
Classification Number: G25
DOI: 10.13266/j.issn.0252-3116.2019.07.014

The advent of the big data era has transformed data into a ubiquitous social presence as an intelligence resource containing significant value [1]. Ordered data collections are regarded as information assets and widely applied. In 2000, the United Kingdom proposed establishing a new scientific research approach—E-science—by leveraging next-generation network technologies and wide-area distributed high-performance computing environments, aiming to enhance scientific collaboration and improve research efficiency through scientific data sharing. With the implementation of E-science projects, data-intensive research has developed. This data-driven research approach has introduced a series of new tasks, including data generation and collection, development of data organization and analysis tools, data encoding, software and hardware support for scientific research, laboratory data management systems, and data openness and sharing [2]. Consequently, scientific data-related research has begun to receive attention.

Data is fundamental to scientific research. As early as the 1980s, France had established a geological research center to provide scientific data for researchers [3]. Foreign scholars such as L. Borgman [4] and A. Gold [5] have analyzed the role of libraries in scientific data management, arguing that libraries should consider their roles and functions in the E-science environment from a strategic planning perspective, and have examined the necessary conditions for libraries to participate in scientific data services, including policies, technologies, and resource reserves. Recognizing the significance of scientific data services for the future development of libraries, the Association of College and Research Libraries Research Planning and Review Committee has identified "research data services (RDS)" as an important direction for library development [6]. After 2010, domestic scholars began to focus on issues related to scientific data in libraries. For instance, Li Dandan and Wu Zhenxin [7] summarized research data management services; Zhang Mengxia and Gu Liping [8] conducted a summary analysis of data curation policies; Qian Jinlin and Liu Guifeng [9] explored hotspots in foreign research data management; and Chen Yuanyuan and Ke Ping [10] reviewed research data services in university libraries. These studies either focused on specific aspects of library scientific data research or covered relatively short time spans in their reviews.

To better understand the current state of domestic library scientific data research, this paper, based on clarifying relevant fundamental concepts, conducts bibliometric and thematic analyses of domestic library scientific data research to provide references for future studies in this field.

2. Concept Definitions

2.1 Data-Intensive Scientific Research

In 2007, Turing Award winner Jim Gray proposed the "fourth paradigm" of scientific research, describing it as data-intensive computing that integrates theory, experiment, and simulation based on data examination [11]. Since then, data-intensive research has gained recognition. Researchers have characterized data-intensive research from different perspectives. First, it is considered a data-driven research approach. Zhang Kaiyong [12] described it as research that does not need to start from scratch but rather achieves its goals through the analysis, organization, parsing, understanding, and utilization of research data. Yao Xiaohua [13] characterized it as scientific research that is no longer limited to the traditional model of theory-hypothesis-experiment-verification, but instead discovers new viewpoints, knowledge, and patterns directly from massive data. Second, it is supported by big data and related technologies, demonstrating dependence on information infrastructure and data processing technologies. Chen Ming [14] viewed data-intensive research as the result of strong intervention by big data technology in various fields of scientific research, with data and information computer technology being the two most important elements. Finally, data-intensive research is based on certain theoretical innovations. Jia Xiangtong [15] argued that data-intensive research takes empiricist scientific methodology as its theoretical foundation, focusing on empirical data itself, and uses data acquisition and analysis technologies to generate theories from data, thereby breaking through the traditional research approach of presetting theories before seeking verification.

Synthesizing current research, data-intensive research exhibits new characteristics in theoretical foundations, research methods, and research support. Its concept should encompass three basic elements: (1) massive data as the starting point of research; (2) big data technology and information technology as technical support; and (3) improved research efficiency as the goal, demonstrating theoretical breakthroughs.

2.2 Scientific Data (Research Data)

The Organisation for Economic Co-operation and Development (OECD), in its "Principles and Guidelines for Access to Publicly Funded Research Data," describes scientific data as data derived from factual records of scientific research that the research community or individual researchers consider useful for research results, including experimental values and images [16]. Current definitions of scientific data include: (1) various types of data generated by human scientific and technological activities and data systematically processed according to certain scientific and technological needs [17]; (2) original basic data reflecting the nature, characteristics, and changing patterns of the objective world obtained through scientific and technological activities or other activities, and data systematically processed and organized to meet the needs of scientific and technological activities [18]; and (3) specifically referring to digitized scientific research data [19].

In terms of research data types, these mainly include: factual, observational, and experimental data [20]; traditional structured, semi-structured, and unstructured data [21]; and process data, semi-finished products, and post-research outcomes [22]. Regarding characteristics, they primarily include basic and original data [23], resource data [24], and strategic data [25]. A summary of relevant descriptions of scientific data concepts is provided in Table 1 .

Based on the context of data-intensive research, this paper comprehensively summarizes that the concept of scientific data should contain two basic elements: (1) data throughout the entire process of human scientific and technological activities (original data, process data, and result data); and (2) data that has been digitized or can be digitized.

Table 1. Summary of Descriptions Related to Scientific Data Concepts - Types: Factual, observational, and experimental; Digital and non-digital; Structured, semi-structured, and unstructured; Process, semi-finished, and result data; Observational, experimental, computational, and reference data; Observational, computational, experimental, and recorded data, etc. - **Characteristics:** Basic; Original; Valuable; Diverse; Disciplinary; Dynamic; Shareable; Continuous; Exchangeable; Associated with literature, etc.

2.3 Data Curation

In 2001, the UK Joint Information Systems Committee (JISC) proposed the concept of data curation, describing it as the maintenance, preservation, and value-added of digital scientific data throughout its entire lifecycle [26]. The Graduate School of Library and Information Science at the University of Illinois described it as proactive, continuous data management throughout the lifecycle of scholarship, science, and education. The Japanese Industrial Standards Committee defined it as managing and improving data from its generation, ensuring continuous supplementation and updating of datasets to enable their discovery, utilization, and reuse at any time [27]. Zhang Ying et al. [27] comprehensively analyzed domestic and foreign definitions and defined it as the process of effectively collecting, storing, processing, and applying scientific data using computer hardware and software technologies.

Domestic scholars' discussions on the concept of scientific data curation have mainly focused on three aspects: implementation subjects, behaviors, and objects. Shen Tingting and Lu Zhiguo [28] argued that implementation subjects include research funding management departments, data service centers, and specialized data curation centers. Chen Yaping and Wu Shufen [29] discussed the feasibility of university libraries implementing scientific data curation. Regarding curation behaviors, Zhang Qiuyan [30] viewed it as purposeful selection and identification of scientific data in a certain field, realizing its value through organization, storage, updating, maintenance, retrieval, and utilization. Yang Helin [31] considered it as continuous supplementation and updating of dynamic data. Wu Minqi [32] regarded it as standardized preservation and maintenance of scientific data. Regarding curation objects, Shen Tingting and Lu Zhiguo [28] limited it to digitally stored scientific data, while Huang Wen et al. [33] considered it as scientific digital information. Some library and information science researchers limited it to scientific data in library collections and databases [34].

A comprehensive analysis reveals that the concept of scientific data curation includes the following elements: (1) diversified implementation subjects, which can be undertaken by existing departments or newly established specialized departments; (2) implementation behaviors covering the entire lifecycle of scientific data, representing a dynamic process based on data utilization and value-added preservation; and (3) implementation objects being scientific data stored in digital form or that can be digitized.

---

3. Literature Survey

3.1 Sample Selection and Methods

To comprehensively obtain literature on domestic library scientific data research, this study searched Chinese databases using CNKI, Wanfang Database, and VIP Database as data sources. The search paths included "subject" (CNKI), "title or keywords" (Wanfang), and "title or keywords" (VIP), with literature sources limited to library and information science research. The search terms "research data," "scientific data," "study data," and "technical data" were used as "subjects (CNKI)" or "title or keywords (Wanfang and VIP)," with a cutoff date of May 1, 2018. After deduplication and removal of irrelevant literature, 495 articles were obtained for analysis. This paper focuses on analyzing the annual trends in total literature volume and keyword frequency.

3.2 Temporal Distribution of Literature

Analyzing the temporal distribution of literature aims to understand the developmental trajectory and trends of domestic library scientific data research. The author conducted statistical analysis of the publication years of the obtained literature, with results shown in Figure 1 [Figure 1: see original paper].

Since 2000, domestic library science research has addressed scientific data issues, with some scholars viewing scientific data as a strategic resource and studying policies related to scientific data open access [35]. From 2000 to 2006, only 7 relevant articles were published domestically, averaging 1 per year. This period can be considered the embryonic stage of domestic library scientific data research. From 2007, when the fourth research paradigm was proposed, to 2010, when Europe and America began focusing on library scientific data services, domestic library and information science published 21 relevant articles over four years, averaging 5.2 per year. This stage can be considered the development period. After 2011, domestic library scientific data research entered a period of high activity, with 467 relevant articles published from 2011 to June 2018, averaging 58.3 per year. Based on research intensity and focus in different periods, this paper divides domestic library scientific data research into three stages, with characteristics shown in Table 2 .

Table 2. Analysis of Stage Characteristics in Domestic Library Scientific Data Research | Period | Stage Characteristics | Main Research Content | |——|———————-|————————| | 2000-2006 | Embryonic Stage | Technical-level data management, scientific data sharing, etc. | | 2007-2010 | Development Stage | Scientific data sharing policies, open access, data licensing, retrieval platform design and implementation, data security, library roles, data sharing principles, policies and regulations, etc. | | 2011-Present | High Activity Stage | Data lifecycle theory, data management plans, research data management models, library roles and missions, research data management infrastructure, talent cultivation, research data standards and norms, national policies, library data service models, research data governance, research data management evaluation, research data copyright protection, data literacy, etc. |

3.3 Keyword Frequency Statistics

Keyword statistics primarily scan research hotspots from a thematic perspective. From the 495 selected articles, 2,030 keywords were obtained, averaging 4.1 per article. Using Excel for statistical processing, keywords with frequency $ $5 were identified as high-frequency words, with distribution shown in Table 3 .

Table 3. Distribution of High-Frequency Keywords in Domestic Library Scientific Data Research High-frequency keywords include: university library, research data management, scientific data management, research data management, data-intensive research, research data services, data lifecycle, data management plan, embedded services, institutional repository, university library, digital library, E-research, scientific data sharing, data literacy education, lifecycle theory, scientific data services, data management services, data repository, digital object identifier, academic library, research library, social

science data, data curation, E-science.

High-frequency keyword statistics reveal several basic characteristics of current domestic library scientific data research: (1) A series of concepts centered on "data" have emerged. For example, the term "research data" has been translated by domestic researchers as "scientific data" (155 occurrences), "research data" (63 occurrences), and "study data" (17 occurrences), with minimal difference in conceptual elaboration. Similar situations exist in translating "data curation." (2) Drawing on foreign research and practical experience is an important pathway for current domestic library scientific data research, as evidenced by analyses of relevant projects in Western countries such as the UK (13 occurrences), USA (10 occurrences), and Australia (7 occurrences), while domestic case studies are relatively scarce. (3) Connections between library scientific data research and other library research areas are evident. The appearance of high-frequency words such as data sharing (41 occurrences), data literacy (22 occurrences), and open access (14 occurrences) demonstrates absorption and reference to existing research findings.

These characteristics show that although current domestic library scientific data issues have received widespread attention, research still has considerable room for expansion, such as further clarifying basic concepts and theories and summarizing localized practical experience. A comprehensive and in-depth understanding of domestic research dynamics, analysis of existing problems, and future research trends are of important theoretical and practical significance for promoting scientific data research in China's libraries.

---

4. Research Topic Analysis

Library scientific data research is an emerging field in library and information science. Based on bibliometric analysis of research literature and following the basic principles of combining theoretical and practical research with general research and library service characteristics, this paper constructs an analytical framework from three aspects: scientific data basic theory, related practices, and library scientific data services to better understand the content and characteristics of current domestic library scientific data research.

4.1 Theoretical Research on Scientific Data

Researchers' discussions on scientific data-related theories focus on data lifecycle theory, small science-large science theory, and data governance theory.

4.1.1 Data Lifecycle Theory Lifecycle theory originated in the energy field in the 1960s and was later widely applied in politics, economics, environment, technology, society, and other fields. The scientific data lifecycle refers to the entire process from creation and initial storage to eventual deletion when obsolete. The significance of this theory for scientific data research lies in describing the basic process of scientific data curation and providing guidance for model construction. Li Hang [36] argued that the data lifecycle represents the process of data processing, storage, and management in specific environments to achieve value mining and utilization. The research cycle is the scientific data lifecycle, and by correlating the research cycle with the data lifecycle, researchers can understand scientists' data needs, thereby determining the basic framework for scientific data management and constructing models. Li Zheng [37] defined and explained the entire process of scientific data management based on this theory and provided a basic framework for scientific data management activities. Meanwhile, researchers have extended this theory in combination with scientific activity characteristics, emphasizing that data curation should consider the duality of the scientific data lifecycle: the data's own lifecycle and the co-lifecycle generated after data association with specific fields [38].

4.1.2 Small Science-Large Science Theory The small science-large science theory was proposed by Professor Derek de Solla Price, a science history expert at Yale University, in the 1960s, depicting a "Pricean vision" of structural and dynamic changes in scientific work [39]. One of its important contributions is revealing the impact of organizational differences on research activities. Zhao Yanzhi [40] proposed establishing a "small science" data curation mechanism embedded in research activities based on the characteristics of small science in research scale and project scope. Li Lirui and Deng Zhonghua [41] emphasized the impact of researcher personalization characteristics on small science data and analyzed the relationship between small science data and scientific data as a whole. Yang Ping and Tian Ye [42] investigated small science data management tools and analyzed obstacles and solutions for small science data sharing.

4.1.3 Data Governance Theory Data governance first emerged in enterprise management and was later absorbed by social governance, library science, and other research fields, with its connotation and scope expanded. Gu Liping, in "Data Governance: Development Opportunities for Library Undertakings," proposed that data-driven approaches have become a new route for scientific progress, making various researchers new knowledge service objects. The data governance ecosystem requires libraries to implement policies for data acquisition, sharing, and reuse, with data librarians playing important roles [43]. Data governance theory proposes that data is an asset and clarifies the diversification of management subjects and the rights-responsibilities relationships therein. Researchers have extended discussions on organizational management from the perspective of data as an asset. For example, Zhou Ting and Zhan Qingdong [44] argued that data is an organization's core asset, and the setting of data management authority and responsibility is actually control over data assets. Liu Guifeng et al. [45] viewed data governance as the allocation of responsibilities and decision-making rights within organizations. Ren Yazhong

[46] distinguished data governance from data management, arguing that data management emphasizes ensuring data integrity through data management systems and storage devices, while data governance focuses on control over data collections and rights allocation in data utilization.

4.2 Practical Research on Scientific Data Management

Practical issues in scientific data management are a key focus for researchers. Scholars have analyzed issues such as the formulation of scientific data management policies and plans, the design of management models, and the construction of management platforms based on domestic and foreign cases.

4.2.1 Formulation of Scientific Data Management Policies and Plans

Policies provide the behavioral basis and guarantee for smooth implementation of scientific data management. Research on scientific data management policies involves policy-making subjects, content composition, and formulation strategies. Wei Yue and Liu Guifeng [47] argued that governments, research institutions, and universities are the main bodies for formulating scientific data management policies. He Qingfang [48] believed that governments, research institutions, universities, and publishing institutions can all serve as policy makers. Dong Kun and Gu Liping [49] argued that educational institutions can also undertake corresponding functions. Regarding policy content composition, Liu Xiwen and Xiong Rui [50] believed it mainly includes data open funding, data quality control, data legal protection, data preservation, and data sharing and utilization. Guo Chunxia [51] categorized it into data management planning policies, data collection and processing policies, data storage policies, and data utilization policies. Zhou Xiaoyan and Zai Bingxin [52] divided it into guarantee elements (time guarantee, applicable subject guarantee, responsible subject guarantee, and superior policy guarantee) and subject elements (general elements and research data management elements). Regarding policy formulation strategies, Ding Pei [53] suggested choosing to expand existing policy or guideline systems or independently formulate new policies, emphasizing conciseness and clarity in content detail, and focusing on the role of decision-making departments such as academic committees or university councils.

A Data Management Plan (DMP) refers to the description and documentation of data management covering the entire scientific research process. Its formulation is influenced by the selected management model and also represents the concretization of the management model's implementation. While policy formulation describes the macro framework of scientific data management, management plans focus on specific implementation of management behaviors. Domestic researchers have analyzed the functions, content, and development tools of scientific data management plans. Yang Shujuan and Chen Jiacui [54] argued that DMP functions to provide norms for the creation, preservation, archiving, and utilization of research data, thereby ensuring consistency, accuracy, and traceability of long-term preservation and utilization of scientific data. Wang

Kai et al. [55] divided it into eight components: administrative data, data collection, documentation and metadata, ethics and legal permissions, storage and backup, selection and preservation, data sharing, and responsibilities and resources. Li Xiangyang et al. [56] believed it should include descriptions related to data selection, identification, storage, preservation, sharing, and dissemination. Researchers have also introduced DMP creation tools and applications such as DMPtool [57] and DMPonline [58].

4.2.2 Establishment of Scientific Data Management Models The establishment of scientific data management models is based on theories such as the data lifecycle and reflects the basic characteristics of their disciplines. The data lifecycle model is the most basic scientific data management model. In addition, researchers have analyzed the establishment of models for research data service needs identification, scientific data description, and management maturity. Li Weimian and Cui Yuhong [59] categorized current research data lifecycle models into four types: chain, matrix, ring, and hierarchical, and analyzed the system structure, key process areas, and key practices of the Scientific Data Management Capability Maturity Model (SDM-CMM). This model divides management maturity into five levels and five key process areas with corresponding goals, with each process area having four common characteristics: execution commitments, execution capabilities, execution activities, and process evaluation. Chi Yuzhuo and Wang Yanfei [60] analyzed four projects—DAF, DCPs, AIDA, and CARDIO—as examples of demand identification models for research data services in UK and US universities. These four projects involve data asset frameworks, data curation infrastructure documents, institutional digital asset assessment, and collaborative evaluation of research data infrastructure and objectives, including three elements: organization, technology, and resources, as well as five scoring levels: acknowledge, act, consolidate, institutionalize, and externalize. Wang Fengyang et al. [61] proposed a "W7+R3" scientific data provenance description model based on provenance technology to ensure effective management of scientific data from the source. W7 includes What (events occurring within the data lifecycle), Why (reasons for events), Where (locations of events), When (time or period of events), Who (people or organizations involved in events), Which (tools or software applied in events), and How (actions of events). R3 includes Reference (reference information for events), Result (results and causes of events), and Remark (relevant notes for events). Dang Hongli and Tan Haibing [62] introduced the Data Management Maturity (DMM) model into the library field, proposing the construction of a library data management and service capability maturity model and introducing its practical application.

4.2.3 Construction of Scientific Data Management Platforms Researchers have analyzed the construction of scientific data management platforms based on domestic and foreign project implementations. Relevant literature reveals the following characteristics of current platform construction: (1) Development of information technology provides technical support for platform construction, with software platforms such as Dspace, Dataverse, Fedora, Nesstar, and DigitalCommons supporting scientific data management platform construction, among which Dataverse and Fedora are most widely applied. (2) Platform construction is based on different purposes and focuses. For example, Li Hongqin et al. [63] built a university dissertation linked data publishing system using the Apache Marmotta platform based on LOD (linking open data) concepts. Li Pan et al. [64] introduced W3C's universal metadata standard DCAT and constructed a government open data platform based on Drupal. Four US universities including Stanford built four research data sharing service platforms (SDR, DC, OIAD, HD) based on Fedora and Dataverse systems [65]. Harvard University developed the scientific data publishing platform DR-NTU (Data) using Dataverse [66]. (3) Diversified implementation subjects. Based on existing cases, scientific data management platform construction subjects include universities [67], government agencies [68], university libraries [65], and industry data centers [69]. (4) Involvement of multiple disciplines and fields, covering natural sciences and social sciences such as biology [70], environmental science [71], and social sciences [72].

4.3 Research on Library Research Data Services (RDS)

Researchers have investigated the implementation of scientific data management services in foreign university libraries and reflected on the status and role of libraries in the data-intensive research paradigm, including the construction of library scientific data service systems, data literacy education, and intellectual property services around scientific data.

4.3.1 Investigation of Foreign University Library Scientific Data Management Services Investigation of foreign library scientific data management service implementation is an important part of current domestic research, with characteristics including: (1) **Investigation methods.** Web surveys, literature analysis, and case analysis are the main methods used, primarily analyzing the implementation of RDS in foreign university libraries through their websites or relevant research literature. Examples include Cui Yingji et al.'s survey of research data management services in US university libraries using web surveys and literature analysis [73]; Wang Manrong et al.'s web survey of research data management services in four Singaporean public universities [74]; and Si Li and Zeng Yueliang's web survey of research support services in university libraries ranked in the top 100 world university rankings in 2017 [75]. (2) **Investigation content.** Content covers user needs, service content, service policies, service platforms, staffing, and data literacy education. Examples include Cheng Xifeng et al.'s survey of user needs for RDM value-added services [76]; Su Min et al.'s investigation of specific measures for research data management in eight foreign universities in the ship and ocean engineering field [77]; and Dong Fang's survey of scientific data management and sharing practices in the US library community [78].

4.3.2 Construction of Library Scientific Data Service Systems Current research literature mainly involves two aspects: service content and implementation paths. (1) **Service system content.** Researchers have analyzed the composition of library scientific data service content systems. Li Xiaohui [79] believed that library scientific data services mainly include technical support, research data organization methods, research data service methods, user information literacy education based on research data, and relevant institutional and personnel configurations. Xiong Wenlong and Li Ruilai [80] believed it includes a data management policy guarantee system, data resource construction system, data processing system, data storage system, data service system, user management system, and human resources system. (2) **Service paths.** Researchers have analyzed implementation paths for library scientific data services. Hu Xuehuan and Qu Baoqiang [81], using Purdue University Library's scientific data management services as an example, believed that libraries can provide services through establishing institutional research repositories, assisting researchers in developing data management plans, metadata processing and preservation, providing data management technology-related training, conducting data management guidance and education, data collection and preservation, data consulting services, and digital object identifier (DOI) services. Zhang Shasha et al. [82] introduced five approaches used by US university libraries: establishing thematic websites, providing FAQs, publicizing consulting emails, establishing reference desks, and holding seminars. Huang Honghua and Han Qiuming [83] introduced six UK university library service methods: providing policy consultation, assisting in developing data management plans, organizing relevant training, building data management platforms, personnel support, service promotion, and open access services.

4.3.3 Library Data Literacy Education Research on data literacy and data literacy education is an important aspect of current domestic library scientific data research, including: (1) **Data literacy.** Researchers have discussed the connotation of data literacy. Zhang Jingbo [84] believed that data literacy generally refers to the abilities researchers should possess in scientific data collection, organization, processing, analysis, and sharing, as well as the ethics and behavioral norms generally followed in these processes. Shen Tingting [85] believed that data literacy can be divided into basic data skills, data acquisition ability, data evaluation ability, data management ability, data usage ability, and data expression ability. Hu Hui and Wu Ming [86] constructed a core content system for data literacy embedded in research workflows and data lifecycles based on the general process of the data lifecycle and mapped a researcher data literacy competency framework. (2) **Data literacy education.** This includes content and methods of library data literacy education. Cai Hongqi [87] believed that data literacy education is an important responsibility of current university libraries and proposed characteristics and models of data literacy education in university libraries oriented toward disciplines. Hu Hui and Wu Ming [88] summarized UK and US university library data literacy education models from five aspects: teaching objectives, teaching objects, teaching approaches, teaching content, and teaching evaluation. Zhang Qun and Liu Yumin [89] built a "5W" three-dimensional model for university library scientific data literacy education systems based on systems theory and synergy theory, discussing it from five aspects: participating subjects, intervention stages, driving factors, educational content, and implementation methods.

4.3.4 Intellectual Property Issues of Scientific Data Researchers treat scientific data as a knowledge asset and discuss issues of scientific data rights allocation and authorization in scientific data sharing: (1) **Scientific data rights allocation.** Researchers believe that constructing a research data rights management framework is an important content of library scientific data management services. Gu Liping [90] believed that research data rights are key to research data management, involving six aspects: stakeholder rights, different disciplinary management methods, different levels of management policies, usage and citation rights, storage and dissemination rights, and public sharing rights. Fei Ruoyun [91] believed that library scientific data management services need to analyze and sort out the relevant rights of scientific data creators, providers, users, and publishers, and designed a research framework for rights issues in library research data management services. Zou Zhongcai et al. [92] divided scientific data management into three stages: planning, storage processing, and publication and reuse, and designed a research framework for rights issues in library research data management services based on rights allocation characteristics in different stages. (2) **Scientific data sharing authorization.** Data authorization is a key issue in scientific data sharing. Researchers have discussed related issues. Liu Runda et al. [93] believed that authorization is the key to resolving the contradiction between scientific data protection and sharing effects, which can be divided into completely free sharing, conditional free sharing, cost-recovery sharing, and paid sharing. Si Li et al. [94] believed that differences in scientific data types and sources affect data authorization and discussed authorization methods for different types and sources of scientific data from the perspectives of data use, rights subjects, and confidentiality levels. Wang Shu et al. [95] believed that scientific datasets have different intellectual property protection methods and standards at the collective and individual levels and discussed open data licensing authorization models and compatibility in combination with intellectual property policies.

---

5. Summary and Outlook

5.1 Analysis of Current Characteristics in Domestic Library Scientific Data Research

Through bibliometric and thematic analysis, this paper has sorted out the main components of current domestic library scientific data research. Based on basic concept analysis with emphasis on scientific data management and library scientific data service practices, this paper identifies the following characteristics:

5.1.1 High Sensitivity to Scientific Data Issues Based on Disciplinary Characteristics Bibliometric statistics show that since the beginning of the 21st century, China's library and information science research has begun to pay attention to scientific data issues. While analyzing scientific data as a component of literature and information, researchers have sensed the important value of scientific data in future scientific research and conducted systematic studies. In 2007, with the proposal of the fourth research paradigm, the impact of data-driven approaches on research activities attracted attention from more disciplines, but library and information science remained the main force in scientific data research. Bibliometric statistics show that library and information science has become the main discipline for domestic scientific data research, ranking first among all disciplines in terms of published literature volume.

5.1.2 Gradually Deepening Research Content, Reflecting Integration of Data Science and Disciplinary Development Domestic library scientific data research overall shows the basic characteristics of gradually deepening research questions and integration with disciplinary development, reflected in both theoretical and practical research. From a practical perspective, providing services around scientific data has become an important direction for current library service transformation. Since 2006, researchers have begun to pay attention to the relationship between scientific data sharing and specialized libraries [96], making scientific data sharing an important research direction for library scientific data services. Subsequently, libraries have expanded service content in areas such as data curation, disciplinary scientific data services, data literacy education, knowledge services, and intellectual property services, influencing library staff quality cultivation, position setting, management systems, and strategic planning. In theoretical research, researchers have absorbed relevant data management theories such as data lifecycle theory, small science-large science theory, and data governance theory, enriching the theoretical foundation of library scientific data research, while combining data science theory with disciplinary development [97] to positively impact the discipline.

5.1.3 Emphasis on Absorbing Foreign Experience with Weak Localization Research Foreign-related research and practice are important theoretical and practical sources for current domestic library scientific data research, a characteristic evident in both theoretical and practical research. Scientific data-related research originated abroad. In terms of practice, there remains a significant gap between China and foreign countries in the implementation of data-intensive research, scientific data curation, and library scientific data service activities. Therefore, domestic library scientific data research takes foreign investigation and analysis as an important research pathway, with related discussions developed accordingly, reflecting weak localization research.

5.2 Research Prospects

As China's scientific research informatization level improves and the national big data strategy is implemented, the importance of scientific data for research activities is further strengthened. In January 2018, the national "Scientific Data Management Measures" were issued, and a number of influential research data centers began construction. Facing issues of data governance and security, intellectual property protection of data, and sharing and collaboration around scientific data, all countries have strengthened the standardization and institutionalization of scientific data management. Against this background where data-driven characteristics of scientific research and social development are increasingly evident, it has become an inevitable choice for library and information science research to strengthen attention to scientific data-related issues and seek its own contributions. Based on the overall state of current domestic library scientific data research and development trends in foreign library scientific data research, this paper suggests that future research in this field could focus on the following aspects:

5.2.1 Enrich Research Methods and Strengthen Systematic Empirical Research Grasping the characteristics of data-intensive research and mastering the laws of scientific data management practice are basic requirements for improving the effectiveness of library scientific data research, with relevant theoretical testing, policy framework formulation, and service settings developed accordingly. Regarding current research, methods such as web surveys, questionnaires, and interviews have been applied to some extent in domestic and foreign scientific data management and university scientific data service research, but systematic empirical research is relatively scarce, and overall progress in current domestic and foreign scientific data-related practices is insufficiently understood. This is reflected in two aspects: first, domestic and foreign investigations focus on limited typical cases through web and literature surveys; second, the systematic nature and depth of investigations are insufficient. In contrast, foreign-related research pays more attention to systematicness and depth, such as Y. Kim and M. Adler's survey of 361 social scientists' data sharing intentions to analyze scientists' data sharing behaviors [98].

5.2.2 Actively Conduct Interdisciplinary Research Scientific data research carries obvious technical labels, which to some extent affects the depth of library scientific data research. Related cases show that universities or university libraries providing scientific data services both domestically and abroad have strong technical capabilities, and the construction of scientific data management platforms and service implementation place high demands on librarians' qualities, reflected in both technical literacy and data literacy. At the same time, related project content mostly has interdisciplinary characteristics. Although user analysis, literature analysis, information literacy, metadata, and information resource construction and sharing are traditional research methods and content that give library and information science research a certain foundation when facing scientific data issues, there are still deficiencies in theoretical construction and method selection, resulting in insufficient disciplinary characteristics in scientific data research. Therefore, library scientific data research should strengthen interdisciplinary cooperation, discovering the unique value of library scientific data research and promoting disciplinary development through joint exploration of scientific data issues with other disciplines. In terms of form, this can be achieved through joint project applications and libraries providing services for other disciplines' scientific data research. In terms of content, it is reflected in seeking connections between scientific data research and existing library research content, such as the "Data Storage" project conducted by Johns Hopkins University Library in the USA, which integrated overall planning for scientific data management and long-term preservation of library scientific data [99].

5.2.3 Conduct Basic Research Based on Disciplinary Characteristics
With the development of current data-intensive research, some basic issues beyond the technical field have begun to emerge, such as the formulation of scientific data management policies, data literacy of relevant personnel, and intellectual property issues of scientific data. In 2014, the UK Joint Information Systems Committee (JISC) promulgated "The Value and Impact of Data Sharing and Curation," emphasizing the economic benefits of scientific data sharing [100], reflecting the economic attributes of scientific data management. Perfecting top-level design for scientific data management at the national level, standardizing the basic literacy and professional qualifications of relevant practitioners, and strengthening intellectual property protection of scientific data have become foundational tasks for the future development of data-intensive research. Library and information science research has a good foundation in these areas. University libraries, specialized libraries, and other types of research libraries are important literature resource centers for research activities and important subjects in the construction of national scientific data management systems. Strengthening related policy research is an important task for libraries. Meanwhile, libraries' experience in information literacy education and intellectual property services is of important help for libraries to conduct data literacy research and scientific data intellectual property research. Library scientific data research can focus on these areas, leveraging disciplinary characteristics to better promote the development of scientific data research in the discipline.

References

[1] Peng Zhihui. Data: The research object of information science in the big data environment [J]. Information Science, 2017(2): 123-131.
[2] Lang Yangqin, Kong Lihua. The fourth paradigm of scientific research: An introduction to Jim Gray's report "e-Science: A revolution in scientific research

models" [J]. e-Science Technology & Application, 2010(7): 92-94.
[3] Bu Zhanling. France establishes uranium geology research center to provide scientific data for explorers [J]. Foreign Nuclear News, 1981(4): 30.
[4] BORGMAN L. Research data: Who will share what, with whom, when, and why? [EB/OL]. [2018-06-14]. http://works.bepress.com/borgman/238.
[5] GOLD A. Cyberinfrastructure, data, and libraries. Part 2: Libraries and the data challenge: Roles and actions for libraries [J]. D-Lib Magazine, 2007, 13(9/10): 1-10.
[6] Chang Wei. E-Science and the transformation of literature and information services [J]. Library and Information Service, 2005, 49(3): 27-30.
[7] Li Dandan, Wu Zhenxin. Analysis of research data management services [J]. Library Science Research, 2012(9): 54-59.
[8] Zhang Mengxia, Gu Liping. Review of data curation policy research [J]. New Technology of Library and Information Service, 2016(1): 3-10.
[9] Qian Jinlin, Liu Guifeng. Review of foreign research data management research [J]. Information Theory and Practice, 2017(10): 130-134.
[10] Chen Yuanyuan, Ke Ping. Review of research data services in university libraries [J]. Library Work and Study, 2017(10): 16-23.
[11] Zhou Xiaoying. The rise of data-intensive scientific research paradigm and the response of information science [J]. Information and Documentation Services, 2012(2): 5-11.
[12] Zhang Kaiyong. University library scientific data services in the data-intensive scientific environment [J]. Library and Information Service, 2013(4): 93-95.
[13] Yao Xiaohua. Development and reflection of secondary literature services in the data-intensive research environment [J]. Library and Information Service, 2013(5): 103-106.
[14] Chen Ming. Data-intensive scientific research: The fourth paradigm [J]. Computer Education, 2013(5): 29-32.
[15] Jia Xiangtong. The dual logic and problems of the "fourth paradigm" in the big data context [J]. Journal of Jiangsu Administration Institute, 2017(6): 14-20.
[16] DIRK P. OECD principles and guidelines for access to research data from public funding [J]. Data Science Journal, 2007(6): 4-11.
[17] Jiang Hong, Zhong Yongheng. International scientific data sharing research [J]. Modern Intelligence, 2008(11): 56-58.
[18] Li Juan, Liu Dehong, Jiang Hong. Research on the current status of international scientific data sharing [J]. Library Development, 2009(2): 19-21.
[19] Sun Liling. Strategies for constructing the professional competence of research data librarians—Based on the analysis and interpretation of the "Research Data Librarian Competency Framework" [J]. New Century Library, 2017(10): 32-36.
[20] Wei Yue, Liu Guifeng. Analysis of foreign university scientific data management and sharing policies based on the data lifecycle [J]. Information Science, 2017(5): 153-158.
[21] Chen Jianxin. Scientific data services: A new field for library services [J].

Library and Information, 2013(4): 93-95.
[22] Li Dandan, Wu Zhenxin. Research on research data citation [J]. Library Journal, 2013, 32(5): 65-71.
[23] Qiu Junping, He Wenjing. Research on the interaction between scientific data sharing and citation behavior [J]. Information Theory and Practice, 2015(10): 1-5.
[24] Wang Jian, Wu Dingfeng. Exploration of a resource utilization efficiency evaluation framework for scientific data [J]. Modern Intelligence, 2015(8): 3-6.
[25] Chi Yuzhuo, Wang Yanfei. Content analysis framework for scientific data citation in scientific data management [J]. Journal of the China Society for Scientific and Technical Information, 2018(1): 43-51.
[26] Wei Haiyan, Wei Junchao. Investigation and analysis of foreign university data curation projects [J]. Library and Information Service, 2014(12): 38-47.
[27] Zhang Ying, Zhang Zhiping, Liang Bing. Research on scientific data management application models [J]. Information Engineering, 2017, 3(4): 71-77.
[28] Shen Tingting, Lu Zhiguo. Scientific data curation methods in different stages of research projects [J]. Information Science, 2013(5): 42-47.
[29] Chen Yaping, Wu Shufen. Role positioning of university libraries in the data curation environment [J]. Journal of Ningbo Education College, 2015(10): 110-112.
[30] Zhang Qiuyan. Research on university scientific data curation [J]. Information Science, 2013(5): 42-47.
[31] Yang Helin. Data curation: New exploration of university libraries in the United States [J]. Journal of Academic Libraries, 2011(2): 18-21.
[32] Wu Minqi. Digital Curation: An emerging research field in library and information science [J]. Library Journal, 2012(3): 8-12.
[33] Huang Wen, Ma Juntao, Su Ye. Scientific data curation and its application in specialized libraries [J]. Chinese Journal of Medical Library and Information Science, 2014(9): 41-44.
[34] Cui Yuhong. New roles of research libraries in the E-science environment: Scientific data management [J]. Library Journal, 2012(10): 20-23.
[35] Chen Chuanfu, Zeng Ming. The significance of scientific data complete and open access policies and their implications [J]. Library Science Research, 2014(3): 69-72.
[36] Li Hang. Construction of an academic library research data management system based on the data lifecycle model [J]. Journal of Library Science, 2016(12): 19-21.
[37] Li Zheng. Research on a university library research data management model based on data lifecycle theory [J]. Journal of Lanzhou University of Arts and Science (Natural Science Edition), 2017(3): 81-85.
[38] Qian Peng. Analysis of the duality of information lifecycle management: Taking scientific data management as an example [J]. Information Theory and Practice, 2013(3): 11-14.
[39] Merton, Garfield. Preface to "Little Science, Big Science and Beyond" [J]. Studies in Science of Science, 2013(12): 1790-1791.

[40] Zhao Yanzhi. The scientific value and practical dilemmas of "small science" data curation [J]. Information Theory and Practice, 2016(12): 56-59.
[41] Li Lirui, Deng Zhonghua. Research on micro-knowledge services for research user small data (Part 1) [J]. Library and Information, 2016(3): 49-54.
[42] Yang Ping, Tian Ye. Research progress on long-tail data sharing [J]. Library and Information Service, 2014(4): 133-138.
[43] Gu Liping. Data governance: Development opportunities for library undertakings [J]. Journal of Library Science in China, 2016(9): 40-56.
[44] Zhou Ting, Zhan Qingdong. Discussion on data governance of university library alliances: Taking FULink as an example [J]. Information Exploration, 2018(2): 34-42.
[45] Liu Guifeng, Qian Jinlin, Lu Zhangping. Research progress on data governance at home and abroad: Connotation, elements, models, and frameworks [J]. Library and Information Service, 2017, 61(11): 137-144.
[46] Ren Yazhong. From data management to data governance: Transformation of library functions in the big data environment [J]. Journal of Sichuan Library Science, 2017(4): 6-9.
[47] Wei Yue, Liu Guifeng. Investigation and enlightenment of UK university research data management policy content [J]. Library and Information Research, 2016(4): 35-44.
[48] He Qingfang. Investigation and analysis of foreign scientific data management policies [J]. Shanghai Library and Information Work Research, 2016(2): 9-13.
[49] Dong Kun, Gu Liping. Research on the policy framework of scientific data opening in several countries [J]. China Science and Technology Resources Review, 2016(5): 51-57.
[50] Liu Xiwen, Xiong Rui. Analysis of characteristics of foreign scientific data open access policies [J]. Information Theory and Practice, 2009(9): 5-9.
[51] Guo Chunxia. Research on scientific data management and sharing policies in research institutions [J]. Information Journal, 2015(8): 147-151.
[52] Zhou Xiaoyan, Zai Bingxin. Research on the formulation of research data management policies in Australian universities [J]. Library Development, 2017(2): 63-70.
[53] Ding Pei. Research on foreign university research data management policies [J]. Library Tribune, 2014(5): 99-106.
[54] Yang Shujuan, Chen Jiacui. Research result dissemination and sharing: Overview of data management plans for UK and US national funding projects [J]. Information Journal, 2012, 31(12): 176-179.
[55] Wang Kai, Peng Jie, Qu Baoqiang. Comparative study of foreign data management plan service tools [J]. Information Journal, 2014, 33(12): 203-206.
[56] Li Xiangyang, Gu Liping, Wang Yanbing. Investigation and analysis of data management plan policies of foreign research funding agencies [J]. Information and Documentation Services, 2016(1): 62-67.
[57] Wu Hairu. Research and reflection on the California Digital Library data management plan tool [J]. New Century Library, 2015(5): 69-72.
[58] Wang Pu. Policies, content, and tools for data management plan formula-

tion in the UK and USA [J]. Library and Information, 2015(3): 103-109.
[59] Li Weimian, Cui Yuhong. Research data lifecycle models and their application in service evaluation [J]. Information Theory and Practice, 2015(9): 38-41.
[60] Chi Yuzhuo, Wang Yanfei. Characteristics and implications of foreign university research data service demand identification models [J]. Library and Information Service, 2016(2): 37-41.
[61] Wang Fengyang, Xu Quanjun, Liu Feng, et al. Scientific data provenance description model and specification design [J]. Research and Practice on e-Science and Technology, 2017, 8(1): 27-34.
[62] Dang Hongli, Tan Haibing. Data management maturity model based on DMM and its application in service evaluation [J]. Library Theory and Practice, 2017(7): 98-103.
[63] Li Hongqin, Zhai Jun, Yuan Changfeng, et al. Research on building a readable and writable linked data platform based on Apache Marmotta [J]. Library Theory and Practice, 2017(7): 98-103.
[64] Li Pan, Zhai Jun, Chen Yan. Construction of a government open data platform based on Drupal [J]. Modern Intelligence, 2016(8): 37-43.
[65] Zhang Shanshan, Wang Zheng. Case study of overseas typical research data sharing service platforms based on open-source software [J]. Library and Information Service, 2017, 61(3): 79-86.
[66] Wang Dandan. User testing research on scientific data publishing platforms [J]. Information and Documentation Services, 2017(6): 56-61.
[67] Zhu Ling, Nie Hua, Cui Haiyuan, et al. Construction of Peking University Open Research Data Platform: Exploration and practice [J]. Library and Information Service, 2016, 60(2): 44-51.
[68] Huang Guobin, Qu Yajie, Wang Shu. Data management function analysis of UKDA and ICPSR social science data publishing platforms [J]. Library and Information Service, 2017, 61(21): 40-48.
[69] Li Yi, Wu Dan. Investigation of open medical scientific data platforms [J]. Library and Information Service, 2015, 59(18): 24-29.
[70] Hong Zhengguo, Xiang Ying. Building a university scientific data management platform based on Dspace: Taking the scorpion species and toxin database as an example [J]. Library and Information Service, 2013, 57(3): 39-44.
[71] Jiang Ying. Development status and content analysis of UK environmental field data publishing platforms [J]. Digital Library Forum, 2018(2): 52-59.
[72] Zhang Jilong, Yin Shenqin, Zhang Yong, et al. Sharing and services of social science data: Taking Fudan University Social Science Data Sharing Platform as an example [J]. Journal of Academic Libraries, 2015(1): 74-79.
[73] Cui Yingji, Shen Hongjie, Xia Ye, et al. Investigation and enlightenment of research data management services in US university libraries [J]. Modern Intelligence, 2018(7): 122-127.
[74] Wang Manrong, Liu Guifeng, Liu Qiong. Investigation and enlightenment of research data management services in Singapore university libraries [J]. Library Science Research, 2018(5): 64-71, 22.

[75] Si Li, Zeng Yueliang. Investigation and analysis of research support services in world-class university libraries [J]. Library and Information Service, 2018, 62(4): 30-41.
[76] Cheng Xiufeng, Wang Xuejie, Xia Lixin. Research on value-added services in research data management systems [J]. Information Science, 2018(10): 77-83.
[77] Su Min, Xu Shuhui, Huang Yanjuan, et al. Preliminary exploration of research data management in foreign university libraries: Taking ship and ocean engineering as an example [J]. Digital Library Forum, 2017(4): 62-67.
[78] Dong Fang. Practice and enlightenment of scientific data management and sharing in the US library community [J]. Information and Documentation Services, 2017(5): 107-112.
[79] Li Xiaohui. Discussion on library research data management and service models [J]. Journal of Library Science in China, 2011(9): 46-52.
[80] Xiong Wenlong, Li Ruilai. Research on library data services based on scientific data management [J]. Library and Information Service, 2014, 58(11): 48-53.
[81] Hu Xuehuan, Qu Baoqiang. Scientific data management services of Purdue University Library in the USA [J]. China Science and Technology Resources Review, 2015(11): 59-64.
[82] Zhang Shasha, Huang Guobin, Di Hongyang. Research and thinking on research data management services in US university libraries [J]. Library Journal, 2016(7): 59-66.
[83] Huang Honghua, Han Qiuming. Research on research data management services in UK university libraries [J]. Library Theory and Practice, 2017(8): 73-77.
[84] Zhang Jingbo. Data literacy education in the big data era [J]. Science, 2013, 65(4): 29-32.
[85] Shen Tingting. Data literacy and its impact on scientific data management [J]. Library Tribune, 2015(1): 68-73.
[86] Hu Hui, Wu Ming. Research on data literacy competency embedded in research workflows and data lifecycles [J]. Library and Information, 2016(4): 125-137.
[87] Cai Hongqi. Research on discipline-oriented data literacy education in university libraries [J]. Library and Information, 2016(3): 138-141.
[88] Hu Hui, Wu Ming. Research on best practices of data literacy education in foreign libraries and their implications [J]. Modern Intelligence, 2016(8): 66-78.
[89] Zhang Qun, Liu Yumin. Research on the construction of a scientific data literacy education system model for university libraries [J]. Journal of Academic Libraries, 2016(1): 96-102.
[90] Gu Liping. Basic framework for scientific data rights analysis [J]. Library and Information Knowledge, 2014(1): 34-51.
[91] Fei Ruoyun. Preliminary exploration of rights issues in the implementation of phased research data management services in libraries [J]. Shandong Social Sciences, 2016(6): 615-617.
[92] Zou Zhongcai. Preliminary exploration of rights issues in the implementation of phased research data management services in libraries [J]. Information

Theory and Practice, 2016(1): 64-69.
[93] Liu Runda, Sun Jiulin, Liao Shunbao. Preliminary exploration of data authorization issues in scientific data sharing [J]. Information Journal, 2010(12): 15-18.
[94] Si Li, Jia Huan, Xing Wenming. Research on copyright protection issues and countermeasures for scientific data [J]. Library and Information, 2015(4): 118-122.
[95] Wang Shu, Wang Hong, Song Xiaodan. Research on intellectual property protection and licensing mechanisms for research data [J]. Library Tribune, 2016(4): 65-71.
[96] Wei Dongyuan. The role that specialized libraries should play in scientific data sharing [J]. Library Tribune, 2006(12): 319-321.
[97] Ba Zhichao, Li Gang, Zhou Liqin, et al. Data science and its impact on the transformation of information science [J]. Journal of the China Society for Scientific and Technical Information, 2018(7): 653-667.
[98] KIM Y, ADLER M. Social scientists' data sharing behaviors: Investigating the roles of individual motivations, institutional pressures, and data repositories [J]. International journal of information management, 2015(35): 408-418.
[99] UC Libraries: Transforming the strategic plan [EB/OL]. [2018-10-14]. http://www.libraries.uc.edu/content/dam/libraries/about/docs/strategicplan.pdf.
[100] The value and impact of data sharing and curation [EB/OL]. [2018-03-09]. https://www.jisc.ac.uk/search.

Author Contributions:
Chai Huiming: Responsible for paper design and writing;
Zhang Libin: Responsible for paper structure formulation and revision;
Zhao Yajie: Responsible for data collation and statistics.

A Review of Research on Research Data of Domestic Libraries

Chai Huiming[1], Zhang Libin[2], Zhao Yajie[2]
[1] Library of Tianjin Conservatory of Music, Tianjin 300171
[2] Nankai University Library, Tianjin 300071

Abstract: [Purpose/significance] This article sorted out and analyzed the current research situation of research data of domestic library, summarized the overall research characteristics and made recommendations for future research on this issue to make research better serve our scientific research and promote the transformation of library services. [Method/process] This paper adopted the research method of combination of literature measurement and subject analysis, focuses on subject analysis. [Result/conclusion] Three basic characteristics of the current domestic research include: (1) Emphasis on scientific data. (2) Deepening the research. (3) Attention to foreign experience and weak localization. Prospects of future research are concluded as follows: (1) Enrich research methods and strengthen empirical research. (2) Promote interdisciplinary research.

(3) Strengthen research on basic issues.

Keywords: data-intensive research; research data; data curation; research data services

---
