Visual Analysis of Domestic Institutional Repository Research from 2000-2018: Postprint

Authors: Chen He, Lin Jing

Date: 2023-10-08T00:00:00+00:00

Abstract

[Purpose / Significance] Currently, IR has developed to a relatively mature stage. Reviewing, analyzing, and summarizing the domestic IR development can provide reference suggestions for the next stage of IR development. [Method / Process] Using Excel, VOSviewer, CiteSpace and other software, along with programming-based data processing methods, this study conducts a visual analysis of 1,362 IR journal papers indexed by CNKI from 2000 to 2018. [Results / Conclusion] Domestic IR publications have experienced a period of rapid growth, a relatively stable period, and a rapid decline period, indicating that domestic IR research has entered a new stage with emerging research points; the high-frequency author groups in domestic IR research are primarily from the Chinese Academy of Sciences system and the university system, with the former's research being more cutting-edge and in-depth than the latter's; journals such as "Library and Information Service", "Modern Information", and "Library Science Research" publish the most IR research papers and are relatively influential journals in the domestic IR research field; domestic IR research hotspots include system construction, resource development, academic communication, knowledge management and services, copyright protection, etc.; possible future research hotspots for IR include research data, research services, auxiliary research decision-making, data value-added services, etc.

Full Text

Visual Analysis of Domestic Institutional Repository Research in China, 2000-2018

Chen He, Lin Jing
Library of Xiamen University, Xiamen 361005

Abstract

[**Purpose/Significance**] Institutional repositories (IR) have reached a relatively mature stage of development. Combing, analyzing, and summarizing the development of domestic IR can provide reference and recommendations for the next phase of IR development. [**Method/Process**] Using Excel, VOSviewer, CiteSpace, and programming-based data processing methods, this study conducts a visual analysis of 1,362 IR-related journal papers indexed by CNKI from 2000 to 2018. [**Result/Conclusion**] The findings indicate that domestic IR publication has experienced three phases: rapid growth, relative stability, and rapid decline, suggesting that current domestic IR research has entered a new stage with emerging research hotspots. High-frequency authors (groups) are primarily concentrated in the Chinese Academy of Sciences system and university systems, with the former conducting more cutting-edge and in-depth research than the latter. *Library and Information Service*, *Journal of Modern Information*, and *Researches in Library Science* are among the journals that have published the most IR research papers and are considered influential in the domestic IR research field. Domestic IR research hotspots include system construction, resource development, academic communication, knowledge management and services, and copyright protection. Future potential research hotspots may include research data, scientific research services, auxiliary research decision-making, and data value-added services.

Keywords: institutional repository; IR; data visualization; knowledge mapping; bibliometrics; research hotspots

1. Introduction

The concept of institutional repository (IR) began to emerge abroad in the late 1990s. In 2001, the Ohio State University Library and the Office of the Chief Information Officer jointly launched the Knowledge Bank Service, allowing employees, faculty, and students to upload their works to the repository for long-term preservation—an early prototype of IR in practical terms [1]. In 2002, Hewlett-Packard collaborated with MIT to develop DSpace, the IR construction software [2], and established MIT's institutional repository DSpace@MIT [3], becoming the first academic institution to formally establish an IR. In 2004, Zhejiang University Library built the first experimental IR in China based on DSpace. In August 2006, Xiamen University Library officially launched and released the Xiamen University Institutional Repository based on DSpace [4], becoming the first university in mainland China to operate a formal IR. The Lanzhou Branch of the National Science Library of the Chinese Academy of Sciences (CAS) began constructing a disciplinary IR platform based on DSpace in 2006, which was subsequently promoted and implemented throughout the CAS system after continuous improvement [5]. Since then, an increasing number of domestic scholars have begun to pay attention to and research IR, making it one

of the discussion topics at major library and information science conferences [6]. Since 2013, the China IR Conference has been held annually to exchange and discuss the latest IR research results and promote IR development in China [7]. To date, IR has been developing in China for nearly 20 years. This study conducts bibliometric and data visualization analysis of domestic IR-related journal papers to examine IR researchers and research hotspots, providing references for IR researchers and offering recommendations for future IR innovation and development.

2. Temporal Distribution of Publications

Statistical analysis of the publication year metadata from the 1,362 retrieved papers was conducted using Excel to create a curve chart of annual publication volume, as shown in Figure 1 [Figure 1: see original paper].

Abroad, IR-related research began to be reported around 2002 [8]. As shown in Figure 1, domestic IR research articles first appeared in 2004 [9], reflecting that Chinese scholars' research closely followed international trends, albeit with a short time lag. From 2004 to 2018, IR-related publications can be roughly divided into three phases:

Phase 1 (2004–2009): This was a period of rapid growth in IR publications, reaching the first peak in 2009 with 133 articles. During this phase, IR was introduced to China as a new academic topic, attracting great attention and research enthusiasm from domestic library and information science scholars, who published numerous academic papers on IR. A review of the papers from this period reveals that they primarily introduced overseas IR research progress and achievements, including the background and conceptual connotations of IR, the functional significance of academic communication, and the implications of IR as an "imported product" for library construction and academic research in China.

Phase 2 (2009–2017): This was a relatively stable period for IR publications. After the accumulation of knowledge from the first phase, domestic library and information science professionals gained a more comprehensive theoretical understanding of IR and began to enter the practical development stage. Publications during this phase involved IR construction practices, resource development exploration, bibliometric analysis, open access policies, disciplinary services, and research management. This phase represented the stage where IR truly took root in China, with diversified research directions and a flourishing development of IR research, with annual publication volumes consistently reaching nearly one hundred articles.

Phase 3 (2018): The total number of publications in 2018 showed a significant decline compared to previous years (the original search was conducted in February; when this paper was revised in June 2019, the number of 2018 publications had increased to 88). The trend suggests that the decline may continue. This phenomenon intuitively indicates that journal papers containing "institutional

repository" and related concepts in their titles, keywords, or abstracts began to decrease. However, does this mean that domestic IR research has stagnated or progressed slowly? The authors searched CNKI using "research data"—a current hot topic in IR—as the search term, retrieving 78 articles in 2018 with this term in their titles. Most of these papers did not contain "institutional repository" or similar terms in the relevant fields, yet they still fall within the scope of IR. Therefore, domestic IR research did not stagnate in 2018; rather, papers using "institutional repository" and related concepts as themes began to decrease, indicating that domestic IR research has entered a new stage with emerging hotspots. The term "institutional repository" alone can no longer satisfy the requirement for retrieving all papers in this field. These new research hotspots will be further analyzed in the "Research Hotspot Analysis" section below.

3. Analysis of Major Publishing Journals

The 1,362 papers were published in 224 journals. Based on the "80/20 rule," the top 20% (45 journals) were selected for statistical analysis, resulting in the proportional diagram shown in Figure 2 [Figure 2: see original paper].

As shown in Figure 2, ten journals—*Library and Information Service*, *Journal of Modern Information*, *Researches in Library Science*, *Information Studies: Theory & Application*, *Information Research*, *New Technology of Library and Information Service* (renamed *Data Analysis and Knowledge Discovery* in 2017), *Journal of Intelligence*, *Sci-Tech Information Development & Economy*, *Journal of Library and Information Science in Agriculture*, and *Library Journal*—accounted for 50% of all IR publications. This indicates that domestic IR researchers submit most of their papers to these journals. Based on these statistics, these journals can be considered highly influential in the domestic IR research field. Users who need to understand specific domestic IR research developments should focus on IR papers published in these journals.

4. Analysis of Paper Authors

4.1 High-Frequency Authors Statistics and ranking were conducted for authors of the 1,362 papers. A total of 21 authors published eight or more papers, which can be considered high-frequency authors in domestic IR research, as shown in Figure 3 [Figure 3: see original paper].

The author with the highest publication frequency is Zhu Zhongming, with 34 papers, ranking first and more than doubling the output of the second-ranked author, making him an ultra-high-frequency author in domestic IR research. In terms of institutional affiliation, authors from the CAS system account for approximately half of the high-frequency authors, while the remaining high-frequency authors are primarily from university systems. This reflects that the CAS system is more active in IR research than university systems, and their research is also more systematic, in-depth, and cutting-edge.

4.2 Author Collaboration Network The author field from the original search results was processed using Python programming and converted into a ris file, which was then used to create an author collaboration network map using the visualization software VOSviewer (with a minimum publication threshold of three papers per author), as shown in Figure 4 [Figure 4: see original paper].

Figure 4 shows four research teams with more than five members: (1) **Zhu Zhongming's research team**, whose members are primarily from the Lanzhou Branch of the National Science Library of CAS. Their publications reveal that this team is responsible for the technical development and maintenance of the CAS IR system. With funding from multiple projects, they completed the development of a new IR system—CSpace—based on the DSpace prototype, tailored to the actual needs of CAS, and promoted its use throughout the academy. According to user needs, the team has continuously developed, updated, and maintained the system and has begun to promote it to university systems outside CAS [10]. This team conducts the most in-depth, comprehensive, and cutting-edge IR research in China and is an authoritative team in the field. (2) **Gu Liping's research team**, whose members are mainly from the National Science Library of CAS. Their publications show that they and Zhu Zhongming's team belong to the same CAS institutional knowledge team, complementing each other, but the former focuses more on the promotion and service of the CAS IR system, open access policy research, and auxiliary system support. (3) **Nie Hua's research team**, whose members are primarily from Peking University Library. Their publications reveal that Peking University Library was involved in IR construction at an early stage and served as the main responsible unit for the Calis-IR project, undertaking IR construction and promotion for CALIS member libraries and currently promoting the construction of the China Academic Library & Information System (CALIS) IR Alliance [11-12]. (4) **Li Guojun's research team**, whose members are mainly from Beijing University of Science and Technology Library. They have primarily discussed IR system construction, digital resource development, and research on literature citation and bibliometric services based on IR, using the Beijing University of Science and Technology IR as a foundation [13].

Additionally, Figure 4 shows that Sun Tan and Xiao Ming serve as bridging figures in the author collaboration network, connecting inter-institutional IR research collaborations among CAS, Beijing Normal University, Beijing University of Posts and Telecommunications, and Beijing University of Science and Technology.

5. Research Hotspot Analysis

Since keywords are not standardized subject terms, it is necessary to perform consolidation before statistical analysis. For example, "open access" was standardized to "open access," various terms like "institutional repository," "institutional archive," "institutional storage," etc., were unified as "institutional repository," and "university" was standardized to "college and university."

5.1 High-Frequency Keywords Through word frequency statistics of the standardized keywords, a list of high-frequency keywords was obtained, as shown in Table 1 (top 30 only).

As shown in Table 1, "institutional repository" has the highest frequency, far exceeding other keywords. The next most frequent keywords are "open access," "library," and "college and university," which also appear much more frequently than other terms. IR emerged in the context of the rapid development of the open access movement and represents a specific practice of open access. Discussions of open access often involve IR, and discussions of IR frequently mention open access-related content and background. Therefore, "open access" has become the second ultra-high-frequency keyword. The keywords "library" and "college and university" indicate that domestic IR research is primarily conducted by libraries, especially academic libraries. An important mission of IR is the long-term preservation of institutional knowledge outputs, which aligns with the library's mission to preserve human knowledge. It can be said that libraries' active research, construction, and promotion of IR services is both a professional and a historical mission. DSpace, as open-source software for building IR systems, has made tremendous contributions to the widespread adoption and development of IR worldwide. Domestic scholars often begin their IR research by studying DSpace, which is why DSpace also appears among the high-frequency keywords. Other high-frequency keywords will be analyzed in conjunction with the keyword co-occurrence network below.

5.2 Keyword Co-occurrence Network Following the author collaboration network approach, VOSviewer was used to create a keyword co-occurrence network map (see Figure 5 [Figure 5: see original paper]). Since "institutional repository," "open access," "library," and "college and university" are ultra-high-frequency keywords that would obscure the co-occurrence relationships of other keywords in the network, they were excluded from the co-occurrence network map. The parameter setting required a minimum keyword frequency of eight occurrences.

Based on Table 1 and Figure 5, combined with a comprehensive analysis of specific literature, several hotspot research areas in domestic IR can be identified:

(1) IR System Construction Research. This hotspot emerged primarily in the early stages of IR research. Multiple software systems are available for building IR, among which DSpace has attracted countless IR enthusiasts due to its open-source code, free availability, rich functionality, and community support. It enabled the transition of IR from theory to practical implementation. Published literature covers the functional architecture of DSpace, comparative studies with other IR construction software, and DSpace localization and specific application practices. Co-occurring keywords include data model, pattern, metadata, interoperability, linked data, and open-source software.

(2) IR Resource Development Research. The foundation of IR services

lies in the collection and management of specific resources. Institutional academic outputs take various forms, including but not limited to journal articles, conference papers, theses, presentations, patents, software, videos, and scientific data. How to effectively collect, organize, and arrange these academic outputs has been a persistent hotspot in domestic IR research. Related keywords include content development, quality control, academic resources, open access resources, intellectual property, self-archiving, service model, strategy, bibliometrics, construction model, and sustainable development.

(3) IR's Role in Promoting Academic Communication Research. With the development of information technology and network technology, as well as changing concepts and methods of academic research activities, the traditional time-consuming, labor-intensive, and costly academic communication model no longer suits the trend of free and rapid academic exchange. The open access concept allows users to quickly and barrier-free access required academic resources through public networks, changing the traditional academic communication model [14]. As a product of the open access movement, IR has achieved the function of free and rapid academic exchange, attracting significant attention and recognition from researchers and general users. Keywords co-occurring with academic communication include knowledge management, information resources, resource sharing, IR alliance, academic IR, DSpace, pattern, and Japan.

(4) IR Knowledge Management and Services Research. The essence of IR is to effectively organize and manage institutional knowledge outputs and make these knowledge products openly available on the network to achieve free academic exchange. To provide IR services, libraries need to not only collect and accumulate original academic resources but also collaborate with relevant institutional organizations, such as university research management departments and personnel management departments, to effectively organize and manage IR content to provide efficient services for different researchers or entities [15]. This represents a higher-level research on IR. Keywords related to IR knowledge management and services include knowledge management, knowledge services, research management, research outputs, data management, big data, resource sharing, subject librarians, linked data, IR alliance, and sustainable development.

(5) Copyright Protection Research. IR emerged in the context of the open access movement, aiming to provide a platform for free academic exchange and barrier-free access to academic resources for users. However, this does not mean that academic resources can be used arbitrarily without protection. How to protect the copyright interests of works and authors while collecting institutional academic outputs and providing external services, and simultaneously offering maximum convenience for users, is also a research hotspot in domestic IR. Keywords related to copyright protection include intellectual property, copyright, self-archiving, quality control, knowledge sharing, fair use, and digital library.

Overall, the boundaries between these research hotspots are not particularly distinct but rather permeate and blend with each other. The research hotspots

are relatively evenly distributed, with no particularly prominent single hotspot.

5.3 Keyword Timeline View Using CiteSpace (5.3.R4.8.31.2018), a timeline view analysis was performed on keyword nodes from metadata downloaded from CNKI in Refworks format, resulting in Figure 6 [Figure 6: see original paper]. To avoid ultra-high-frequency keywords obscuring other keywords, "institutional repository," "open access," "library," and "college and university" were excluded from the analysis.

As shown in Figure 6, CiteSpace automatically categorized keywords into nine clusters using its built-in algorithm. Among them, the timeline for cluster #0 (institutional repository) originated earliest and most completely demonstrates the development process of IR in China, beginning with the introduction of DSpace open-source software for building IR systems in 2004, followed by DSpace localization to adapt to domestic application environments, and learning from the IR construction experiences of Taiwan region universities, which share the same language and culture and were involved in IR construction earlier. With the promotion and popularization of IR, topics such as IR construction strategies, construction models, and service models were subsequently discussed by scholars. Throughout this process, advanced IR construction practices from abroad (such as Japan and the United States) continued to be introduced and referenced.

Examining the timelines of all nine clusters, only clusters #3 and #4 extend to 2018, indicating that these two categories represent current research hotspots. Combined with newly emerging keywords on other timelines, new research tendencies in domestic IR can be summarized, which may 预示 new research hotspots:

(1) Research Data. Among newly emerging keywords, "data" appears most frequently, indicating that domestic IR research has begun to focus on the field of research data. Research data falls within the scope of IR resource development. In past IR resource development studies, the main focus was on explicit academic resources with "output" characteristics, such as journal articles, conference papers, presentations, and theses. Research data in scientific activities were often considered "intermediate products" or supplementary materials to academic outputs and were frequently neglected and abandoned after research activities concluded. However, as scientific activities enter the fourth paradigm and computational science advances, particularly in the context of big data research, data reuse has been recognized and valued, prompting attention to the preservation and reuse of valuable data from research activities [16].

(2) Research Services. IR was born for scientific research and will ultimately serve it. Scientific activities are dynamic and continuous processes that constantly change and adjust with the advancement of open science and changes in socioeconomic environments. Similarly, IR serving scientific research will continuously self-adjust and improve, representing a persistent hotspot in future

IR research.

(3) Auxiliary Research Decision-Making. With the development of big data science and the continuous improvement of IR content and functionality, IR will play an increasingly important role as a "think tank" in assisting institutional research decision-making. Of course, auxiliary research decision-making is not based solely on a single institutional repository but may be built upon "IR alliances" across multiple institutions. IR participation in auxiliary decision-making is a continuous goal pursued by domestic IR researchers and builders, requiring long-term effort and commitment.

(4) Data Value-Added Services. After years of development, IR has encompassed almost all publicly published academic outputs from institutions and some unpublished academic outputs, and has begun to involve research data. Therefore, IR will possess advantages in academic data and, in addition to providing conventional basic services, will offer data value-added services [17] in the wave of "double first-class" university construction.

Additionally, CSpace, as a localized IR construction software system launched by the Lanzhou Branch of the National Science Library of CAS, has been well promoted and used within the CAS system. With continuous updates and improvements and expanding user groups, future publications about CSpace may maintain a certain volume.

6. Analysis and Discussion

Through the visual analysis of 1,362 papers from different perspectives, the basic state of IR research in China can be understood: domestic IR research and construction have achieved remarkable accomplishments and are now at a critical juncture. Based on the current status of domestic IR research and construction, the authors propose four recommendations for future IR development.

6.1 Continuously Advance Institutional Repository Construction According to statistics [17], there are 325 IRs built or under construction in mainland China's universities and research institutes, with resource types including journal articles, theses, conference papers, books, patents, research reports, presentations, and experimental data. In other words, many Chinese universities and research institutes are currently building large-scale repositories of academic outputs—academic "big data." As the nation increasingly values data and has elevated big data to a national basic strategic resource, considering it the "oil" of the information society [18], continuing to advance IR construction is of major "strategic" significance with broad development and application prospects.

On the other hand, as a platform for continuously collecting and managing intellectual outputs generated by institutional teaching and research activities, IR construction should not cease as long as teaching and research activities continue. IR construction should be continuously advanced through policy and funding support from multiple dimensions.

6.2 Integrate and Optimize to Become Part of the "National Basic Strategic Resource" Currently, domestic IR construction has achieved remarkable results, building a large number of IR systems primarily containing institutional academic outputs, with resources reaching a considerable scale. Two major IR resource alliances have emerged with resource sharing as their main purpose: the "Knowledge Repository of the National Science Library, Chinese Academy of Sciences" and the "China Academic Library & Information System (CALIS) Institutional Repository Alliance." These resources should be positioned as components of "national basic strategic resources" in line with national development trends. However, these resources currently cannot achieve true interconnection and sharing, remaining scattered across different educational institutions and research institutes, with resource entities still controlled by individual institutions.

Resources only generate value when they flow. Therefore, it is necessary to integrate and optimize current IR resources to enable their interconnection and sharing. However, thoroughly integrating and optimizing national IR resources inevitably involves various mechanism conflicts and interest adjustments between different educational and research institutions, and this process cannot be completed by a single institution. Instead, it requires national-level reforms to existing IR policies, systems, and mechanisms to coordinate and balance multiple interests. IR builders need to liberate themselves from the mindset of "fighting 各自的为战" and broaden their thinking and vision for IR resource development and utilization. Only through thorough integration and optimization can IR resources become true "national basic strategic resources."

6.3 Coordinate and Collaborate to Promote In-Depth IR Development As revealed in the hotspot analysis, a considerable number of authors from different institutions are conducting identical or similar IR research topics, and the two major IR research groups (CAS group and university group) are also undertaking similar IR research and IR alliance construction. Due to the relative dispersion of IR researchers or groups and the lack of unified planning and coordination, IR research hotspots are widespread but fail to form true "hotspots," while causing phenomena of redundant investment, repeated research, duplicate outputs, repetitive labor, and redundant construction, wasting personnel, energy, and even financial resources. These problems are inherent in the early spontaneous, 粗放式, and "anarchic" development model of IR.

IR development today must reform this previous development model through comprehensive coordination, division of labor, and collaboration among national IR research groups and individuals to avoid redundant research. Various IR research groups and individuals should conduct targeted research on hotspot issues in IR development trends based on their respective characteristics and advantages, such as research data hotspot studies and research assistance hotspot studies, forming IR hotspot research with certain distinctive features. Simultaneously, at the national level, a group of capable IR research teams should

be organized to conduct pre-research and 攻关性 studies on hotspot issues in IR development trends. For example, as IR resources become "national basic strategic resources," exploring how to use big data technologies and methods for their in-depth development and utilization can give Chinese IR research distinctive characteristics while forming a tiered IR research team.

Through comprehensive planning and coordinated coordination, breakthroughs and innovations in IR research and construction can be achieved economically and efficiently, promoting the in-depth development of domestic IR.

6.4 Promote Construction Through Evaluation to Improve IR Quality and Level As mentioned earlier, early domestic IR construction was relatively 粗放式, lacking guidance and unified planning. The wide variety of IR resources and varying levels of institutional investment in IR construction have resulted in common problems such as relatively 单一的系统功能, inconsistent resource types, and non-standard or incomplete resource cataloging. These issues affect subsequent IR system upgrades and resource development and utilization.

To enable better and faster IR development and facilitate IR resource interconnection and sharing, establishing an IR construction quality evaluation system is an effective means. Literature surveys show that scholars both domestically and internationally have proposed IR quality evaluation indicator systems. R. Serrano-Vicente et al. proposed evaluating Spanish IRs based on criteria related to technology, procedures, content, marketing, and personnel [19]. Guo Qingrong proposed ideas and specific indicators for establishing an IR evaluation system [20]. Guo Yi proposed an IR evaluation system consisting of nine principles and their corresponding performance indicators [21]. Tang Qi proposed an IR evaluation indicator system based on the Balanced Scorecard [22].

However, IR construction is a dynamic process that requires a flexible and dynamic evaluation system. Therefore, from the perspective of national strategic objectives, different evaluation indicator systems should be adopted for different stages of IR construction to conduct timely assessments, ensuring that IR construction moves toward high quality and high levels.

7. Conclusion

As this study's data source is limited to journal papers indexed by CNKI, with search terms being "institutional repository" and related concepts, and does not include other academic research data such as doctoral and master's theses, the research results may deviate from the actual situation to some extent. Additionally, other metadata fields with certain analytical significance (such as funding sources, author affiliations, and references) were not included in this study due to non-standard source data, which should be incorporated in future research.

The authors' recommendations for future IR development are macro-level suggestions proposed from a national perspective, hoping that relevant national

departments will take the lead in coordinating the future development of domestic IR, organizing domestic IR research experts, editors and reviewers of core IR journals, and big data application experts to form an IR advisory decision-making group. Based on summarizing past IR construction experiences, this group would conduct top-level planning and design for future IR development, determine development strategies, and introduce supporting policies. These suggestions may be somewhat idealistic and are intended to spark discussion among domestic IR research colleagues about the future development of IR.

References

[1] Zhong Yuan. Bibliometric analysis of foreign institutional repository research from 2004 to 2013[J]. *Information Science*, 2016, 34(6):71-76.
[2] DURASPACE. DSpace - A Turnkey Institutional Repository Application[EB/OL].[2019-03-30]. https://duraspace.org/dspace/.
[3] MIT Libraries. DSpace@MIT[EB/OL].[2019-03-30]. http://dspace.mit.edu/.
[4] Xiamen University Library. Xiamen University Institutional Repository[EB/OL].[2019-03-30]. https://dspace.xmu.edu.cn/.
[5] Zhu Zhongming, Ma Jianxia, Chang Ning, et al. Research and practice on building disciplinary knowledge repository systems based on DSpace[J]. *New Technology of Library and Information Service*, 2006(7):10-14.
[6] Shanghai Society of Library Science. Previous Advanced Seminars on Digital Library Frontier Issues[EB/OL].[2019-03-30]. http://society.library.sh.cn/node/2827.
[7] China IR Promotion Working Group. Previous Conferences[EB/OL].[2019-03-30]. http://2018chinair.csp.escience.cn/dct/page/70002.
[8] Roy T. Institutional repositories[J]. *Library Journal*, 2002, 127(15): 28-30.
[9] Wu Jianzhong. Library vs. Institutional Repository—Rethinking Library Strategic Development[J]. *Journal of Library Science in China*, 2004(5):7-10.
[10] Zhu Zhongming, Zhang Wangqiang, Liu Wei, et al. Strategies and models for open source software practice in China's digital libraries: A case study of CSpace[EB/OL].[2019-03-30]. http://dlib-oss2014.csp.escience.cn/dct/attach/Y2xiOmNsYjpwZGY6MjU1NzY=.
[11] Nie Hua, Wei Chengfu, Cui Haiyuan. CALIS institutional repository: Construction, promotion, reflection, and prospect[J]. *Journal of Library Science in China*, 2013, 39(2):46-52.
[12] Wei Chengfu, Nie Hua, Cui Haiyuan. Multi-library collaborative development of institutional repositories—A case study of the CALIS institutional repository project[J]. *Journal of Academic Libraries*, 2014, 32(3):69-73.
[13] Li Guojun, Wang Yu, Wang Limei, et al. Research on university institutional repository construction based on metadata—A case study of Beijing University of Science and Technology Institutional Repository[J]. *Journal of Academic Libraries*, 2012, 30(4):55-60.
[14] Li Chunwang. Open access to academic information in the network environment[J]. *Journal of Library Science in China*, 2005(1):33-37.
[15] Wei Tongqi, Fan Guangbing, Weng Shuo, et al. Utilization and value-added services of institutional repository resources in research libraries[EB/OL].

[2019-03-30]. http://ir.psych.ac.cn/handle/311026/27283.
[16] Chen He. Discussion on the construction of scientific data repositories in domestic universities[J]. *Digital Library Forum*, 2017(12):45-51.
[17] Gong Yinong, Zhu Ming. Survey on the current status of institutional repository construction in China[J]. *Digital Library Forum*, 2018(9):20-28.
[18] Yang Bo. Big data becomes a national basic strategic resource, and government data interconnection and sharing promote industrial takeoff[EB/OL].[2019-03-30]. http://www.gov.cn/zhengce/2015-08/20/content_{2916507}.htm.
[19] Serrano-Vicente R, Melero R, Abadal E. Evaluation of Spanish institutional repositories based on criteria related to technology, procedures, content, marketing and personnel[J]. *Data Technologies and Applications*, 2018, 52(3): 384-404.
[20] Guo Qingrong. Construction of an institutional repository self-archiving evaluation system[J]. *Journal of Intelligence*, 2009, 28(7):74-76, 73.
[21] Guo Yi, Ji Ping. Construction of an institutional repository evaluation system[J]. *Library Science Journal*, 2014, 36(2):10-12.
[22] Tang Qi. Construction of an institutional repository evaluation indicator system based on the Balanced Scorecard[J]. *Library Work and Study*, 2019(4):53-58.

Author Contributions:
Chen He: Conceived the research idea, collected and processed data, wrote and revised the manuscript.
Lin Jing: Processed data and created figures, wrote the manuscript.

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