Digital Resource Preservation Assessment Based on Metadata Management: A Postprint Study

Authors: Xu Tiancai, Pan Yuting, Yang Xinya, Luo Li, Sun Rui

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

Abstract

[Purpose/Significance] The extensive management model of digital resources can no longer satisfy the refined information service demands of university libraries. Constructing a novel digital resource management system based on literature metadata repositories, improving metadata-based digital resource procurement agreements and oversight processes, and establishing a multi-dimensional literature assurance evaluation system on this foundation will provide important reference for enhancing information service quality in university libraries.

[Method/Process] This study employs literature research methodology and case analysis to explore development strategies for university libraries in digital resource management and literature assurance evaluation.

[Result/Conclusion] The new paradigm of fine-grained digital resource management provides the foundation for multi-dimensional literature assurance evaluation in university libraries. Only with a systematic and standardized digital resource management and evaluation system can university libraries adapt to the development trends of the digital information era and effectively leverage big data analysis to support the advancement of university teaching and research.

Full Text

Preamble

Volume 63, Issue 2, January 2019

Research on Digital Resource Guarantee Evaluation Based on Metadata Management

Xu Tiancai[1], Pan Yuting[2], Yang Xinya[1], Luo Li[3], Sun Rui[1]
[1]Chongqing University Library, Chongqing 400044
[2]School of Economics and Business Administration, Chongqing University,

Chongqing 400044
[3]Chongqing Medical University Library, Chongqing 400016

Abstract

[**Purpose/Significance**] The extensive management model for digital resources can no longer meet the refined information service needs of university libraries. Constructing a new digital resource management system under a literature metadata warehouse, improving metadata-based digital resource procurement agreements and supervision processes, and establishing a multi-dimensional literature guarantee evaluation system will provide important references for enhancing the quality of university library information services.

[**Method/Process**] Using literature research and case analysis methods, this paper explores the development strategies of university libraries in digital resource management and literature guarantee evaluation.

[**Result/Conclusion**] The new form of fine-grained digital resource management provides the foundation for multi-dimensional literature guarantee evaluation in university libraries. Only with systematic and standardized digital resource management and evaluation systems can libraries adapt to the development trends of the digital information era and truly utilize big data analysis to support the development of university teaching and scientific research.

Classification Number: G251
Keywords: digital resources; metadata; resource management; guarantee evaluation
DOI: 10.13266/j.issn.0252-3116.2019.02.010

Resources have always been the foundation of library operation and development. With environmental changes and technological upgrades, libraries have transformed from "paper resource owners" to "digital resource users," and from "leaders" to "appliers" of resource organization techniques [1]. However, current digital resource management in domestic and foreign university libraries largely remains at the database level, without penetrating into the database to conduct fine-grained management of specific resources.

The primary task of university libraries is to provide adequate resource guarantees for teaching and scientific research construction, followed by providing the most convenient information retrieval services for readers. However, without mastering the specific characteristic information and content of individual resources at the bottom layer of databases, libraries cannot achieve clear, efficient, and accurate management and evaluation of digital resources, nor can they provide the foundation for convenient cross-database one-click retrieval for readers. The numerous and diverse databases not only increase the difficulty of information retrieval for readers but also cause libraries to become simple intermediaries with massive resources but no basic data [2]. The extensive and vague management of digital resources at the database level cannot accurately

reflect the digital resource guarantee strength of university libraries in teaching and scientific research construction, nor can it reveal detailed differences in literature guarantees compared with benchmark institutions and disciplines. Improving the quality of literature guarantee work in university teaching and scientific research construction has important reference value for enhancing the quality of library digital resource management and evaluation work.

2 Metadata Management

2.1 Current Status of Metadata Management

Metadata is a data set that describes and specifies data characteristics, relationships, and corresponding operations [3]. Libraries have consistently paid insufficient attention to metadata during digital resource procurement, with procurement agreements typically lacking relevant content regarding literature metadata. In a survey of 20 databases within the contract period jointly procured by university library digital resource procurement alliances (including Nature, ProQuest Dissertations & Theses, ScienceDirect Online, IEEE/IET Electronic Library, Begell Digital Library, ASCE Online Research Library, etc.), 14 databases made no mention of metadata provision, while only 6 databases indicated that metadata could be provided [4]. The methods of metadata provision are also ambiguous, with no clear specifications regarding whether it involves automatic harvesting, self-downloading by customers, or network transmission by data providers. This current situation of neglecting metadata management must be addressed with urgency.

Represented by Chongqing University Library, domestic university libraries have begun metadata harvesting and organization, receiving active cooperation from publishers such as Elsevier, Wiley, and Springer. Based on Chongqing University Library's experience in metadata harvesting, most data providers are willing to cooperate, though some require prior agreement signing. Electronic book data is easier to obtain than journals, with some journals only providing volume-level rather than individual article-level data. It is evident that database providers have strong protection awareness regarding metadata, particularly making explicit restrictions on metadata dissemination, copying, and transfer to third parties. Metadata harvesting is a long-term, continuously updated task, and ensuring timeliness of harvesting and metadata cleaning are key and challenging points in the management process. After prolonged exploration, Chongqing University Library has achieved certain results, with the harvested and cleaned data volume reaching over 160 million entries (see Figure 1 [Figure 1: see original paper]). Well-known university libraries such as Peking University, Tsinghua University, and Shanghai Jiao Tong University also attach great importance to literature metadata harvesting and management, involving relevant metadata usage in academic evaluation and analysis as well as institutional repositories.

Although libraries still face many difficulties in metadata provision, collection,

and management, the feasibility of literature metadata harvesting and management is unquestionable. Conducting fine-grained metadata management, improving procurement agreement protections, and implementing full-cycle supervision and data collection will be of great significance for library digital resource management and utilization.

2.2 Fine-Grained Metadata Management

Fine-grained management of digital resource metadata refers to refining digital resources in library databases to underlying units such as individual volumes, articles, and figures, and assigning standardized unique collection identifiers to each computational unit according to unified rules. Users will no longer need to learn complex database retrieval rules and can conduct cross-database searches through the library's knowledge discovery system. Libraries can also achieve refined comparison and cross-database management, adjusting procurement content and improving digital resource guarantee strength and utilization through data analysis based on specific usage patterns. On the basis of fine-grained digital resource metadata management, libraries can accurately provide derived literature proactive push services based on user retrieval information, and analyze and manage user access and download data.

By mastering users' usage of digital resources, including access and download counts down to the "article" and "volume" level in databases, and combining this with actual database costs and university discipline construction planning factors, libraries can form analysis reports. Based on these reports, libraries can conduct targeted digital resource addition and reduction work. This benign resource utilization system with effective interaction between libraries and users will undoubtedly better ensure the orderly progress of literature construction work for university teaching and scientific research.

2.3 Procurement Agreement Protection

The current extensive database management model based on "databases" as units is disconnected from users' increasingly refined service demands. Digital resource management with "articles," "volumes," and other units as granularity is a challenge that libraries must face directly and the foundation for creating precise and convenient services for users. In the process of integrating and managing library digital resources, the importance of literature metadata is self-evident. Libraries should require data providers to supply metadata in digital resource procurement agreements and provide detailed descriptions of metadata harvesting and usage rights in procurement agreements, protecting the rights of both data providers and libraries while fully considering libraries' multifaceted needs for digital resource management.

Specifically, procurement agreement clauses should include the following key contents:

(1) **Standards for digital resource metadata provision.** The quality

of metadata will directly affect the implementation of refined management work. Procurement agreements must make clear requirements for metadata that data providers should supply, proposing a set of standards including element content, acceptable or compatible metadata data standards, and field necessity to ensure the accuracy of metadata provided by data providers and reduce the difficulty and workload of data cleaning for libraries in the future [7].

(2) **Methods of digital resource metadata provision.** Procurement agreements should specify the specific methods by which data providers supply metadata, including but not limited to API self-crawling, FTP download, Excel lists, Marc data, XML lists, Access data, etc., while clearly noting the economic costs involved in different methods. Libraries can also parallel-crawl metadata used by users during their access to databases through the digital library webpage, using computers for automated collection.

(3) **Scope and permissions of digital resource metadata usage.** Data providers must guarantee the quantity and quality of metadata provided to libraries, while libraries should also refrain from disseminating, copying, or transferring metadata to third parties to protect data providers' interests. Therefore, libraries and data providers should establish corresponding usage scopes and permissions in procurement agreements to standardize this work.

(4) **Other provisions.** Digital asset procurement is a complex task for both libraries and data providers, not a simple transaction. Preliminary procurement agreements should also include provisions on metadata update frequency, after-sales services to be provided by data providers, handling of violations by both parties, exemption clauses, supplementary agreement terms, etc.

2.4 Full-Cycle Supervision and Data Collection

Digital resources demonstrate different functions and values at various stages of their lifecycle, and libraries need to conduct long-term effective dynamic supervision and data collection of digital resources [8]. The specific process of digital resource lifecycle supervision is shown in Figure 2 [Figure 2: see original paper].

All digital resources renewed and newly purchased by libraries each year should be strictly supervised with detailed documentation archived for subsequent work. During the use of digital resources, libraries should conduct long-term, dynamic supervision of resource access and download counts, regularly recording, summarizing, and analyzing supervision results. Library supervision and maintenance of digital resources is a key link in literature resource guarantee work for university teaching and scientific research construction. The data obtained through supervision will effectively guide the structural optimization of digital

resources and, to a certain extent, analyze the research profiles of various university disciplines, providing references for leadership in discipline construction and planning.

When libraries refine digital resource management to the individual article level, through dynamic supervision, they can regularly filter out expired, invalid, erroneous, and authorization-problematic data, and remove or delete these unusable data for various reasons, ensuring the overall quality of library digital resources.

3 Multi-Dimensional Digital Resource Guarantee Evaluation

Based on metadata management, libraries have clearer control over their digital resources and related usage, which provides important basic data support for subsequent digital resource guarantee evaluation work. Establishing a resource guarantee evaluation system based on literature metadata fine-grained management will be more targeted, using multi-dimensional evaluation criteria to analyze literature data, thereby effectively regulating library resource allocation and enabling libraries to be well-informed and purposeful in their support for university teaching and scientific research construction.

3.1 Evaluation Criteria

When purchasing digital resource types and quantities, libraries generally make acquisitions based on the university's discipline distribution, procurement recommendations from relevant departments, and recommendations from faculties, scholars, and database providers [9], which to a certain extent meets the resource needs of most users. However, with the proposal of building "double first-class" universities, new requirements have been put forward for libraries' digital resource construction capabilities and guarantee levels. Libraries must meet the more refined literature needs of key advantageous disciplines during their construction process, establish a more comprehensive digital resource guarantee system, understand procurement needs from multiple angles and in all directions, and comprehensively evaluate digital resource usage intensity to satisfy the literature information resources needed for key advantageous discipline construction, teaching, and scientific research.

The new evaluation criteria should focus on measuring whether they meet the needs of key advantageous disciplines, including authoritative discipline literature resources, scholar publication resources, derived literature resources, benchmark institution literature, research hotspots, and other series of requirements, as detailed in Table 2 .

3.2 Evaluation Strategies

Digital library resource construction requires scientific formulation of resource construction plans, focusing on six aspects: resource construction principles,

construction methods, construction work content, construction strategies, construction funds, and construction management. Among them, construction strategies require libraries to strive to expand resource volume and improve resource guarantee levels and supply capabilities [10]. Faced with limited budgets and exponentially increasing digital resources, university libraries must conduct strict evaluations, comprehensive considerations, and select the best options, while fully considering university discipline development plans and reader usage evaluations to achieve timely elimination and renewal. The specific process is shown in Figure 3 [Figure 3: see original paper].

In summary, libraries need to form a complete multi-dimensional evaluation system to assess their literature guarantee work in university teaching and scientific research construction. Using multi-dimensional evaluation criteria for refined analysis of existing resources to obtain effective information on library resource guarantees can guide libraries in adjusting their current resource structure. At the same time, based on analysis results, libraries can clarify literature needs for university education and scientific research, providing direction for procurement.

(1) Key Advantageous Discipline Literature Guarantee Evaluation. The literature resource guarantee for key advantageous disciplines must focus on both breadth and depth construction [11]. For key advantageous disciplines in universities, implement discipline clustering, emphasize interdisciplinary approaches, and strive for comprehensive coverage to meet the breadth requirements of digital resource construction. As various key advantageous discipline majors continue to refine and deepen, they must align with all majors related to key advantageous disciplines, maintain resource professionalism, and achieve completeness of resource types to realize depth construction of digital resources. University libraries should clearly identify their own key advantageous disciplines and corresponding authoritative digital resources in their fields, prioritize the procurement of authoritative domestic and international databases, top-tier journals, and core journals in key advantageous discipline fields, and highlight the support and guarantee strength and importance of digital resources for key advantageous disciplines on the digital library portal homepage, using limited funds for more important discipline digital resource construction work to achieve optimized resource allocation and improve the library digital resource guarantee system.

(2) Scholar Publication Resource Guarantee Evaluation. Scholar publication resource guarantee evaluation refers to the guarantee status of digital resources such as journals and monographs where university scholars publish their academic achievements. Since scholar publications constitute the main component of university academic achievements, and academic research largely has inheritance relationships, it is necessary to ensure the continuity and systematicity of scholar academic resource guarantees for deeper academic exploration and research. University libraries should fully investigate the publication status of their scholars and evaluate the resource coverage of their scholars' publica-

tions in databases. Since scholars have very clear knowledge of their published academic papers, a composite approach of "voluntary submission and active collection" can be used to obtain scholars' academic achievements. Libraries should regularly analyze and process registered and reviewed data, filter out missing journals and databases, and prioritize meeting the digital resource needs of journals and databases where scholars publish high-level and high-impact papers.

(3) Derived Literature Resource Guarantee Evaluation. Academic research often requires tracing origins and investigating literature derived from academic documents, such as references, citation literature, co-citation literature, and co-cited literature, which have important reference and research value for target fields [12]. Therefore, libraries cannot ignore digital resources involved in derived literature of papers in their acquisition and cataloging work and must systematically and comprehensively count resources involved in derived literature. Based on the guarantee of derived literature resources, resources can be guaranteed in order of their closeness to the original literature, sequentially ensuring references, co-citation literature, co-cited literature, secondary references, secondary citation literature, and other digital resources. Evaluating the guarantee status of derived literature resources can help assess the completeness of university library digital resource construction work.

(4) Benchmark Institution Literature Resource Comparative Guarantee Evaluation. Comparative analysis and evaluation of literature resource guarantee strength with benchmark institutions is currently missing from library work content. University libraries should effectively conduct benchmark institution literature resource comparative evaluations, selecting appropriate benchmark institutions that target higher levels than their own discipline development and digital literature resource guarantee levels. Conduct multi-angle evaluation and analysis from multiple levels such as digital library webpage construction and database guarantee strength, continuously narrowing gaps through comparison, clarifying their own missing literature resource types, and actively filling current gaps in digital resources to establish correct resource construction development directions, narrow gaps with benchmark institutions, and enhance the digital literature resource guarantee level supporting key advantageous discipline construction.

(5) Research Hotspot Literature Resource Guarantee Evaluation. How libraries guarantee literature resources for research hotspots and emerging research fields is also important work that cannot be ignored [13]. On one hand, libraries need to understand relevant theme content for national project applications each year in advance, fully investigate new fields and research hotspots involved in project themes, evaluate missing parts in their database resources, and timely acquire and supplement scarce literature information resources. On the other hand, libraries need to cooperate fully with university scientific research management departments, requiring subject librarians to maintain close contact with researchers in various disciplines to ensure libraries can understand the latest research trends and hotspots of various schools,

disciplines, and researchers at the first moment, and timely conduct guarantee evaluation work for digital literature resources in hotspot research fields.

(6) Literature Resource Guarantee Evaluation Based on Linkage Mechanisms. When formulating digital resource acquisition and cataloging strategies, libraries need to comprehensively consider budget constraints, discipline development, opinions of subject librarians and scientific researchers, and comprehensively evaluate existing and missing literature resources to make objective and accurate acquisition decisions. Therefore, establishing literature resource guarantee evaluation based on linkage mechanisms requires literature resource guarantee construction from different professional perspectives: Linkage between subject librarians and scientific researchers, where libraries can deeply understand literature resources needed for discipline development and improve digital resource utilization among discipline professionals through the role of subject librarians; Linkage between libraries and various schools, where libraries can comprehensively consider literature resource retention based on actual needs of each school, ensuring objectivity, authenticity, and reliability of discipline digital resources, and more accurately providing digital resource guarantees for discipline construction and evaluation; Active response to university development plans in emerging key research fields, providing information services in advance and strongly supporting related discipline construction and development; Fine-grained digital resource metadata management provides libraries with the possibility to clearly understand the average usage cost of unit literature (article, chapter, volume), and can also derive the connection between individual articles and related scientific research outputs of the university, thereby judging the effectiveness of digital resource usage and ultimately guiding library procurement decisions.

3.3 Elimination Mechanism

Investigating user satisfaction with digital resource usage is an important aspect of evaluating digital resource guarantees. Libraries should establish a user-demand-oriented digital resource guarantee model [14]. Using big data mining to collect users' daily usage of digital resources and understand their evaluations of current digital resources, libraries should conduct real-time monitoring of purchased digital resource usage while comprehensively evaluating purchase costs and utilization rates, implementing an elimination mechanism for lowest-performing resources to allocate resource funds where they are most needed. This can not only save library resource costs but also improve the overall utilization level of library digital resources.

For database purchases, a trial evaluation mechanism should be adopted to establish a supplementary trial digital resource database ranking. Based on resource usage during the trial period and user evaluation feedback, libraries can form a purchase intention ranking list and make corresponding decisions on whether to purchase databases [15]. By combining resource elimination mechanisms and trial evaluation mechanisms, libraries can replace lowest-performing

digital literature resources with top-ranked trial evaluation resources, ensuring the effectiveness and sustainability of library digital resource guarantees through a benign "metabolism" cycle of digital resources.

3.4 Using Evaluation Results to Regulate Resource Structure

Through refined analysis of existing resources using the above multi-dimensional evaluation criteria to obtain effective information on library resource guarantees, libraries can be guided in adjusting their current resource structure. At the same time, based on analysis results, libraries can clarify literature needs for university education and scientific research, providing direction for procurement. However, most suppliers currently negotiate with libraries using bundled procurement models such as databases, putting libraries at a clear disadvantage. This is mainly because domestic libraries did not clearly define their needs during the initial stage of digital resource construction and blindly expanded resource volume, allowing suppliers to bundle and sell much redundant data. However, as domestic libraries become more rational in digital resource construction, more and more libraries have begun effective commercial negotiations with internationally renowned publishers such as Elsevier, Springer, Wiley, and Wolters Kluwer. University libraries such as Peking University, Shanghai Jiao Tong University, Tsinghua University, Tongji University, and Nanjing University have achieved annual adjustments to resource content at the database level, replacing some journals and books or conducting independent journal selection, though the problem of excessively high single-journal costs still exists.

Library digital resource construction must have clear direction, and this direction must be based on analysis results grounded in large amounts of demand data. Although libraries cannot achieve absolute freedom in current procurement, as long as needs are clearly defined and resource structures are continuously adjusted, even if the scope is limited, resource structures will certainly become more suitable for university construction needs.

References

[1] Cheng Huanwen, Huang Mengqi. Between "Paper Worship" and "Digital Adoption"—The Dilemma and Way Out of Information Resource Construction in University Libraries[J]. Library Forum, 2015(4): 1-8.

[2] Xu Tiancai, Yang Xinya, Peng Wandong. Development Status and Trends of New Generation Library Management Systems—A Summary of the 2016 Ministry of Education University Library Committee Information Technology Application Annual Meeting[J]. Journal of Academic Libraries, 2016(6): 5-9.

[3] Lin Haiqing. Metadata System for Digital Libraries[J]. Journal of Library Science in China, 2000(4): 59-64, 69.

[4] Sun Rui, Yang Xinya, Wei Qunyi, et al. Research on Key Issues in Literature Asset Metadata Warehouse Construction—Taking Chongqing University

Library as an Example[J]. Journal of Academic Libraries, 2018(2): 18-24.

[5] Chongqing University Library. Homepage Hongshen Search[EB/OL]. [2018-06-06]. http://lib.sccm.cn/library.htm?g=1001395129967334.

[6] Sun Rui, Yang Xinya, Liao Wei. Constructing Literature Asset Management Based on Metadata Warehouse[J]. Library Forum, 2017(7): 9-16.

[7] Ding Qiujin, Ma Yuanyan, Li Bohui. Research on Organizational Management Framework in Multi-Source Metadata Integration[J]. Digital Library Forum, 2017(12): 58-62.

[8] Tian Xiaodi, Sun Boyang. Full-Process Management Functions of Electronic Resources in Next-Generation Library Service Platforms—Taking Alma as an Example[J]. Library and Information Service, 2016, 60(17): 65-71.

[9] Jiang Yanbo, Chen Xiangzhu. Research on Digital Resource Construction Issues in Domestic Higher Finance and Economics College Libraries—Based on Survey Analysis of 45 Finance and Economics College Libraries in China[J]. Library and Information Service, 2015, 59(8): 65-71.

[10] Shen Xiaojuan, Zhou Chen, Han Chao. Interpretation of "Digital Library Resource Construction Guidelines"[J]. Journal of Library Science in China, 2011, 37(1): 38-46.

[11] Jiao Kun, Xia Qiuju, Li Zhenglong. Citation Digital Resource Guarantee Analysis of SCI Papers for Disciplines[J]. Journal of Intelligence, 2012, 31(7): 72-75, 90.

[12] Baidu Baike. Significance of Citation Literature[EB/OL]. [2018-03-11]. https://baike.baidu.com/item/%E5%BC%95%E8%AF%81%E6%96%87%E7%8C%AE/1195609?fr=Aladdin.

[13] Jiang Yanbo, Sheng Xiuwen. Research on Regional Literature Resource Guarantee System Construction for Key Disciplines in Local Universities—Taking Jiangxi Changbei University Library Alliance as an Example[J]. Library and Information Service, 2015, 59(11): 89-93.

[14] Li Haoling, Xiao Long, Xu Cheng. Application of User Satisfaction Survey Method in Digital Resource Evaluation[J]. Journal of Academic Libraries, 2007(1): 41-46.

[15] Du Yingqi. Selection and Evaluation Strategies for Trial Digital Resources in University Libraries[J]. Library Construction, 2011(7): 13-15.

Author Contributions

Xu Tiancai: Designed the overall research framework and wrote the paper;
Pan Yuting: Responsible for basic data compilation and chart editing;
Yang Xinya: Responsible for partial content revision and research material provision;

Luo Li: Responsible for partial content revision;
Sun Rui: Responsible for partial content revision.

Abstract: [Purpose/significance] The extensive management model for digital resources can no longer match the refined information service needs of university libraries. Setting up a new digital resource management system under the literature metadata warehouse, improving the metadata-based digital resource procurement agreement and supervision process, and establishing a multi-dimensional literature guarantee evaluation system will provide an important reference for improving the information service quality of university libraries. [Method/process] Using literature research and case analysis, this paper probes into the development strategy of university libraries in digital resource management and literature guarantee evaluation. [Result/conclusion] The new form of fine-grained digital resource management provides the basis for multi-dimensional literature guarantee evaluation of university libraries. Only with systematic and standardized digital resource management and evaluation systems can university libraries adapt to the development trend of the digital information era and truly use big data analysis to promote the development of teaching and scientific research.

Keywords: digital resource; metadata; resource management; security assessment

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