

Exploration and Practice of Institutional Repository Services at Tongji University: Post-print

Authors: Liu Yueru, Wang Congjun, Yu Yuren, Yuan Jinbo

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

Abstract

[Purpose / Significance] By briefly describing the service practice experience of Tongji University Institutional Repository, this paper explores the service modes of institutional repositories.

[Method / Process] Based on the service practice experience of Tongji University Institutional Repository, this paper conducts user group demand analysis, provides a detailed introduction of the characteristic service model based on scientific research information management needs as well as service effectiveness, and summarizes factors that need to be considered in institutional repository construction, including data quality, usage details, and publicity methods.

[Results / Conclusion] This paper discusses future development trends of institutional repositories, such as extended services for novelty search and citation, and expanded services for mobile usage.

Full Text

Introduction

Institutional repositories, developed based on the “open access” concept, provide services for institutions to manage and disseminate their digital resources. Essentially, these repositories serve as long-term preservation platforms for electronic resources, functioning as crucial infrastructure for managing research outcomes, disseminating academic knowledge, and supporting societal innovation. They have increasingly become vital components of institutional knowledge infrastructure and important tools supporting digital research and education [?].

In early institutional repository research, Wu Jianzhong [?] defined institutional repositories as knowledge bases that collect and preserve intellectual resources from single or multiple university communities. P. S. Graham [?] argued that

research libraries need to establish repositories for long-term storage of electronic scholarly resources. C. A. Lynch [?] defined institutional repositories as university-based systems that preserve and provide access to digital scholarly materials. One of the primary functions of early institutional repositories was to archive institutional academic achievements.

Through continuous practice, institutional repositories have gradually expanded their functions and evolved their service philosophies. Yang Helin [?] noted that the growing reliance on research output statistical analysis and research information management systems for research evaluation has created opportunities for libraries to strengthen bibliometric work and institutional repository development, enabling integration and interoperability between institutional repositories and research information systems. In April 2016, COAR (Confederation of Open Access Repositories) launched the “Next Generation Repositories” project, which positions institutional repositories as distributed, globally networked infrastructure for academic exchange platforms [?]. This initiative signifies a shift from mere storage to value-added services, promoting research and innovation while enhancing connectivity and collaboration among institutional repositories [?]. In November 2017, COAR released its first report describing 11 new behaviors, technologies, standards, and protocols—such as social networking and peer review—that facilitate new service development [?].

The traditional functions of institutional repositories do not conflict with these emerging roles; rather, archiving—the traditional function—serves as the foundation for these expanded capabilities. When developing institutional repository services, institutions should adopt need-based approaches tailored to different user groups and requirements. Previous studies have rarely documented the practical experiences of domestic university libraries in developing institutional repository services. This paper draws on Tongji University’s institutional repository service practices to discuss service targets, effectiveness, and models, summarize key considerations for repository construction, and explore future development trends.

2. User Groups of Tongji University’s Institutional Repository

Domestic university institutional repositories have evolved beyond mere knowledge archiving to incorporate additional functions, such as collaborating with institutional functional departments to support research performance data management. Peking University’s institutional repository (PKUIR), as a component of Peking University’s academic exchange ecosystem [?], functions as a research outcome management subsystem within the university’s comprehensive research information management system, offering rich functionalities and various value-added services for scholarly research and exchange. It also enables interoperability with Peking University’s data platform for data interaction and discovery. Tsinghua University’s institutional repository has shifted from IR construction to scholar profile development [?], transitioning from institution-based to

scholar-based management of academic outputs, providing basic academic information search and faculty research profile display services to enhance the repository's influence.

Shanghai University Library [?] has developed its institutional repository based on commercial software, focusing on the university's research and management needs with the goal of "providing basic outcome data for the university and supporting research management decision-making." The repository provides data support for Shanghai University's appointment period assessments, annual evaluations, and mid-level professional title evaluations, with outcome data directly embedded in the personnel information system. Beijing University of Science and Technology's institutional repository [?] ensures timely annual data updates through collaboration with the university's Personnel Department and Research Management Office, addressing the sustainable development challenges of institutional repositories.

International university institutional repositories also offer numerous case studies. For instance, Middle Tennessee State University (MTSU) Library [?] collaborated with the university's Honors College to develop repository content through a three-step process: evaluating existing resources, assigning responsibilities for various resources, and maintaining and expanding resources. This collaborative experience demonstrated that clear division of responsibilities is crucial for library publishing and long-term preservation of institutional academic resources.

Tongji University's institutional repository has developed multiple services tailored to the university's characteristics, summarizing different service models for diverse user needs, including joint construction services, technical services, and data services.

The repository serves three primary user groups: scholars (faculty), academic departments, and functional departments. For scholars, the most fundamental service is the centralized display of personal academic achievements through individual profile pages, showcasing academic accomplishments and expanding scholarly influence. Faculty profile pages display not only basic information such as photos, ORCID numbers, and research directions but also list academic achievements during their tenure at Tongji University. By scanning a QR code on the page, users can quickly access a faculty member's profile, effectively serving as an "electronic business card." Through the institutional repository, faculty can demonstrate their academic "charisma" and expand their influence. For research performance statistics, the repository enables "one-time claim, minimal data entry, universal use," simplifying the process of academic outcome reporting.

For academic departments, the repository provides data-based support, helping department management quickly understand the department's research output overview and trends to inform decision-making. Additionally, Tongji's repository data is relatively comprehensive, meeting departmental research perfor-

mance statistics needs and simplifying the workload of research secretaries.

For functional departments, the repository provides various research performance data tailored to their specific needs, enabling them to quickly grasp the university's research landscape, from disciplinary development trends to high-level paper 收录 status, reflecting the institution's research strength and academic level from multiple perspectives. Some departments prioritize data outcomes, for which the library conducts relevant analysis based on institutional repository data and provides results according to departmental needs.

Overall, services should address the different needs of various user groups by leveraging different strengths of the institutional repository, as illustrated in Figure 1 [Figure 1: see original paper]. The three user groups are not isolated “islands” but are organically connected and interrelated. Faculty participation in claiming and submitting helps “fill gaps” and expand the repository's data coverage. Departments' “top-down” approach of requiring faculty participation and having research secretaries review data ensures data reliability, which in turn encourages functional departments to recognize and endorse the repository—a novel form of promotion. The institutional repository serves faculty, departments, and functional departments while also benefiting from these user groups.

3. Service Models of Tongji University's Institutional Repository Based on Research Information Management Needs

In recent years, universities and research institutions both domestically and internationally have explored the relationship between institutional repositories and research information management systems, achieving certain results in practice, such as at Cornell University, University of St Andrews, Peking University, National Science Library of the Chinese Academy of Sciences, ShanghaiTech University, and Shanghai University [?]. Applying institutional repository paper data to institutional research information management represents one of the current practical application directions for institutional repositories.

Tongji University's institutional repository has developed services related to research information management. Since different user groups have varying needs—some requiring basic paper data for performance allocation, others focusing on technical services to obtain paper data and author contributions, and still others needing processed data based on specific indicators—the repository provides different types of services to meet diverse user requirements. Leveraging data advantages and targeting value-added services, the repository has innovatively developed multiple service models, which can be summarized into three categories, as shown in Figure 2 [Figure 2: see original paper].

3.1 Joint Construction Services

Joint construction services involve the library collaborating with academic departments to build the repository together. This collaboration is not solely the library's responsibility; departments must also participate in data construction. Departments need to “mandate” faculty participation in repository claiming and submission, review faculty actions, and vouch for data reliability.

As shown in Figure 3 [Figure 3: see original paper], the institutional repository 收录 publicly published academic achievements, pushes 疑似 personal 成果 to faculty through algorithmic identification of paper-author relationships, and enables faculty to claim and submit 少量 un 收录 academic 成果. Department research secretaries review claimed and submitted data, with approved data entering the repository's research performance database and being downloaded through the repository backend into departmental assessment systems. In later stages, other data from departmental performance systems are shared back to the repository, creating a virtuous cycle.

Throughout this process, librarians and departments engage in deep collaboration. Librarians not only guide department research secretaries on efficient repository use but also assist in cleaning faculty data, such as confirming common English name abbreviations and habitual paper affiliation addresses—a crucial step for improving matching accuracy between faculty and papers. When encountering disputed data, both parties collaborate to trace data origins and discuss solutions, accumulating special cases that inform subsequent algorithm improvements.

This collaborative model represents a win-win scenario. For faculty, “one-time claim, minimal data entry, universal use” simplifies reporting procedures. For departments, it provides quick access to research performance data and even a stable, free performance platform. For the institutional repository, it supplements data that could not be automatically 收录, ensuring data quality and making the archived data more complete and accurate. Through this collaborative model, a healthy “academic outcome 收录 ecosystem” is formed.

Departments' “top-down” approach of mandating faculty participation and reviewing data ensures reliability, earning recognition from functional departments. Subsequent data sharing further expands the repository's 收录 scope. This service model does not require database technical capabilities; departments can obtain convenient information services directly through the repository backend. It is the most common and easily 推广 able service model.

3.2 Technical Services

3.2.1 Model Overview Technical services involve library data librarians synchronizing institutional repository data to user systems through technical means to achieve real-time synchronization. Faculty claims or submissions in the institutional repository can be synchronized in real time to various assessment

systems of functional departments or academic departments. During service delivery, it is essential to ensure that technical personnel from all parties can read and correctly understand the repository's tables to obtain needed data.

As shown in Figure 4 [Figure 4: see original paper], using Oracle GoldenGate, institutional repository data is synchronized to the real-time data warehouse of the university's Information Office, which is responsible for subsequent data transmission and security, ensuring real-time data interaction between the institutional repository and other university systems. Meanwhile, the Information Office shares existing data with the repository, such as patent and 著作 information, expanding the repository's 收录 scope.

In the technical services model, the Information Office serves as both a collaborative partner and a technical service recipient of the library. The institutional repository database structure is complex, involving multiple outcome types (e.g., papers, 著作) and various claim statuses and methods, with intricate relationships between tables. Library data librarians must explain the database structure to the Information Office, guide users to understand table-field relationships from a library and information science perspective, and assist in writing SQL statements to help the Information Office quickly grasp the repository structure and provide needed data for departments. Through this technical services model, the repository serves multiple functional departments, such as the Development Planning Department and the Personnel Department for talent recruitment and personnel assessment systems.

3.2.2 Practical Case Tongji University's institutional repository began collaborating with the Personnel Department in September 2018 to provide paper data for the personnel system. Through gradual 磨合, as application scenarios evolved and demands increased, the Personnel Department's requirements for data timeliness grew higher, leading to the involvement of a new functional department—the Information Office—in collaborative data services.

Cross-departmental collaboration unfolded in three stages, as shown in Figure 5 [Figure 5: see original paper]:

Stage 1: Data Exchange. The Personnel Department provided previously collected faculty-paper correspondence data (confirmation relationships) to the institutional repository. The library analyzed and parsed these data, reasonably importing them into the repository. Due to incomplete or erroneous early field information, approximately 11% of data could not be utilized. This stage enriched and improved repository data, laying the foundation for subsequent large-scale faculty claiming by reducing future workload through importing existing claim data.

Stage 2: Pre-stored Data. The Personnel Department attempted to interpret the institutional repository database structure, understand table-field relationships, reasonably select needed paper data and fields, and modify relevant codes in the personnel system. In this stage, repository paper data were

pre-stored for use in the personnel assessment system during faculty title reviews. When filling out title review data, faculty did not need to manually enter academic outcome details such as journal names, ISBNs, volumes, issues, or page numbers. By clicking a “load” button, they could directly 调用 data already claimed in the institutional repository, while also being able to modify these data within the personnel system. Data claimed by faculty in the institutional repository were updated to the personnel system in real time, ensuring timeliness. However, information modified by faculty in the personnel system could not be updated back to the institutional repository. This stage primarily targeted faculty undergoing title reviews and served as a “warm-up” phase for utilizing the institutional repository.

Stage 3: Real-time Synchronization. Based on the Information Office’s real-time synchronization warehouse, this stage completely “opened up” the institutional repository database and the Personnel Department’s assessment system, achieving true real-time data synchronization. The implementation involved more faculty—not limited to those applying for titles—and more outcome data—not restricted to outstanding academic achievements. During the year-end summary reporting period, over 2,000 faculty participated in claiming.

Cross-departmental collaboration brought significant convenience to all participating departments. Collaboration with the Personnel Department provided detailed paper data, eliminating the need for faculty to fill in 繁琐 details and earning faculty recognition. For the library’s institutional repository, broader data application increased faculty familiarity with the repository, transforming passive use into active participation.

3.2.3 Model Advantages Collaborating with the Information Office to provide technical services can expand the user base. The Information Office is responsible for real-time data synchronization, ensuring data security during transmission and preventing data leakage by user departments. For the library, the focus can remain on source data quality and stable repository operation. For the Information Office, the collaboration expands its “big data” warehouse with authoritative paper data, enabling more related services. The partnership is not competitive but a friendly strategic collaboration, achieving a “1+1>2” advantage where both parties expand their business scope, providing paper data services not only for the Personnel Department’s various assessment and recruitment systems but also for other functional departments such as the Graduate School’s advisor system and similar real-time synchronization services for some academic departments.

The core of this model lies in providing technical services. The library’s service focus is on SQL-related technical services, primarily targeting users with strong technical capabilities who wish to directly access the institutional repository database.

3.3 Data Services

Data services involve library data librarians processing institutional repository data according to user needs and providing data services to users, as shown in Figure 6 [Figure 6: see original paper]. The library service team organizes requirements based on user objectives and, leveraging knowledge, methods, and metrics from library and information science, helps users achieve their goals. During this process, the team cleans data based on the institutional repository, writes programs to process data, and ultimately delivers expected data for decision support, performance assessment, or other purposes. This model can be understood as the institutional repository indirectly serving users. The service foundation is institutional repository data, with the core lying in data processing. This service model provides personalized services based on different user data needs.

3.4 Summary of Service Models

All three service approaches are based on institutional repository data but differ in service focus and user participation levels:

- **Joint Construction Services:** Require departmental cooperation, with faculty and research secretaries taking initiative to actively claim, submit, and review data. This necessitates departmental “mandates,” such as recognizing institutional repository data for faculty performance assessment and disciplinary evaluation. During service delivery, departments’ key tasks involve promoting participation; they do not need technical support or platform maintenance. The institutional repository provides data and platform. This is the most common and easily 推广 able service model.
- **Technical Services:** Database operation capability is a prerequisite, required for both library data librarians and users. Full database synchronization improves data timeliness and convenience but demands higher technical capabilities from users.
- **Data Services:** Users have lower participation levels, needing only to communicate requirements and clarify objectives with librarians. This service is based on institutional repository data with a focus on data processing.

When serving users, multiple service models can be combined. As shown in Figure 7 [Figure 7: see original paper], when serving the College of Architecture and Urban Planning, the technical services model provided data and database frameworks, while the joint construction services model utilized the institutional repository platform.

Having data is the library’s advantage, but data providers are replaceable. Providing high-quality services based on data is the library’s core value and irreplaceable contribution. For libraries, how to leverage institutional repositories to provide services is an important research question.

4. Service Effectiveness of Tongji University's Institutional Repository

Tongji University's institutional repository began construction in 2016 and officially launched in May 2017. In June 2017, the library began collaborating with the College of Civil Engineering, using the institutional repository to serve the college. After half a year of "running-in," research secretaries understood the repository's functions, and the library understood the college's needs to develop personalized features. Through large-scale faculty claiming and problem resolution, the joint construction services model gradually took shape. Subsequently, the same model was applied to serve the School of Transportation Engineering and the School of Marxism. Through joint construction and technical services models, the institutional repository also served the College of Architecture and Urban Planning.

In June 2018, Tongji University's institutional repository began collaborating with the Information Office to share repository data with the university's big data warehouse. After multiple attempts and explorations, Oracle GoldenGate was ultimately adopted for data synchronization. Following implementation, the library and Information Office jointly served the Development Planning Department, providing research performance data to facilitate its understanding of university-wide academic development trends and disciplinary conditions for decision-making. Relying on the institutional repository, the library served the Quality Management Office, assisting in paper data reporting for "Undergraduate Teaching Basic Status Monitoring Data"; and served the Research Management Department, processing foreign core data to assist performance management.

Expanding the institutional repository's service scope can elevate the library's status within the university, thereby defending its irreplaceability. Meanwhile, university functional departments can simplify data collection work, achieving a multi-win situation.

5. Sustainable Development and Expansion of Institutional Repositories

University libraries are facing transformation, and the gradual rise of institutional repositories presents both challenges and opportunities. Leveraging institutional repositories can reshape the library's position within the university, helping it escape marginalization. The fundamental attribute of institutional repositories is resource preservation. If only this function is considered, the repository operator is replaceable. Libraries should fully leverage their advantages, such as cataloging skills and subject services, utilize institutional repository data, integrate various information services, and provide value-added services to different user groups in a user needs-oriented manner.

5.1 Building Self-Trust: Ensuring Data Quality as the Service Foundation

Institutional repositories must earn the trust of the library itself. Only when the library has confidence in and recognizes the repository's data quality can it promote the repository to other users. Improving data quality is a critical task for libraries. The authors believe that repository data quality should be built on two dimensions: completeness and accuracy.

Completeness encompasses not only the volume of data stored but also the variety of data types archived. Institutional repositories can preserve rich resource types beyond papers, including patents, 著作, software, and grants. In the initial construction phase, the focus can start with journal and conference papers to ensure paper data quality. Once the repository gains recognition and credibility within the university, other outcome types can be gradually added. In practice, achieving 100% 收录 is challenging and requires scholars to voluntarily submit difficult-to-收录 data. The joint construction services model can effectively expand the repository's 收录 scope.

Accuracy refers to the high precision of data pushed to scholars for claiming. Sacrificing a small portion of data to improve push accuracy may cause scholars to distrust the repository and hesitate to use it across multiple scenarios. Conversely, ensuring comprehensive, omission-free data at the cost of significantly reduced accuracy may annoy scholars. When non-personal 成果数量 approaches or exceeds actual 成果数量, scholars may develop a 排斥心理 toward the repository. Therefore, institutional repositories should continuously adjust author disambiguation algorithms to improve push precision. Algorithms can incorporate artificial intelligence and machine learning with multiple judgment dimensions, considering not only author name variations but also habitual affiliation addresses and research directions to achieve precise pushing.

During repository construction, one of the library's responsibilities is data quality control. Excellent data quality can attract users and "recruit" other systems to interface with the institutional repository, gradually making it a data hub for the university and creating multi-win situations with other departments.

5.2 Building Scholar Trust: Focusing on User Needs and Expanding Value-Added Services

Institutional repositories must earn scholars' trust. Among the various user groups [?], scholars constitute the largest group, and their satisfaction is a crucial factor for the repository's robust development. Functional detail optimization can significantly enhance scholars' user experience. By refining service details, improving service quality, and demonstrating repository value, value-added services can be gradually expanded.

When promoting the institutional repository, different strategies should be adopted for different audiences with a positive attitude toward addressing

questions from all sides to achieve good promotional effects. For decision-makers (university leaders, functional department heads, and college leaders), introductions should focus on overall repository functions, emphasizing data quality and credibility. For implementation layers (research secretaries), specific workshops on repository utilization should be conducted. Workshops can raise awareness of the repository's existence and teach research secretaries how to use it for research performance evaluation. For scholars, various forms of promotion should be used to transition faculty from awareness to active use.

Different service objects have different needs, requiring diversified service models. For faculty, the institutional repository is not just a paper claiming platform but also an "official" platform for showcasing academic levels. The repository can provide more value-added services, demonstrating added value through details. For example, faculty scholar cards can allow self-selection of displayed content such as research directions and personal profiles, but academic achievements can only display data approved through the repository, ensuring data reliability while meeting individual customization needs.

For academic departments, the institutional repository can provide different service models based on departmental needs. The joint construction services model requires cooperation from college leadership to "mandate" faculty claiming and submission, with the repository providing high-quality data and a stable platform. The technical services model requires departments to have technical capabilities and database operation skills to obtain more precise data for subsequent performance data construction, with librarians providing technical support. The data services model requires departments only to receive final data results, with lower participation levels and librarian services focusing on data processing based on repository data.

For functional departments, the repository's advantage lies in comprehensive and accurate data, helping them make global judgments and utilize data. Whether through data services or technical services models, user needs must be precisely met.

During service construction, user needs should be emphasized and value-added services provided. Satisfying various user groups that recognize the institutional repository will make them more satisfied with usage details and value-added services, gradually transitioning scholars from "policy-driven" or "performance-driven" use to willing, active use.

5.3 Building Institutional Trust: Active Promotion and Multi-Party Cooperation

Institutional repositories must earn the university's trust. When the university level recognizes and uses the repository in multiple contexts, the repository gains greater value. Through multi-dimensional and multi-directional promotion and publicity, the repository's visibility can be gradually increased while collaborating with various university departments.

6. Extended Functions and Expansion of Institutional Repositories

6.1 Extended Function: Integrating Novelty Search and Citation Services

The increasing demand for literature verification reports in recent years has posed difficulties and challenges for libraries. Utilizing institutional repositories to enable self-service generation of literature verification reports is feasible. Liu Yanmin et al. [?] designed and developed a self-service export function for citation reports based on Cspace, achieving over 95% accuracy in 收录引用 after testing and evaluation.

Tongji University's institutional repository has relatively high-quality paper data, covering basically all fields involved in literature verification services, and maintains paper-scholar correspondence relationships. Based on these two points, integrating the institutional repository into literature verification services is feasible. This integration can both facilitate readers' issuance of verification certificates and improve the accuracy of repository claiming data.

6.2 Extended Service: Providing Mobile Services

In the digital age, information technology has developed rapidly, and mobile hardware such as smartphones has become widely 普及, changing how users access libraries from traditional physical visits to more mobile access [?]. The authors believe that introducing mobile services to institutional repositories is an inevitable trend.

Tongji University's institutional repository has begun developing mobile services in collaboration with the university's Information Office, embedding key repository functions into Tongji University's App "Tongxin Cloud." Using the university's App can both facilitate faculty completion of multiple campus affairs through a single application and reduce the mobile development workload for the repository side while promoting the repository to faculty through the application. Implementation should consider personnel authentication access methods and embed services suitable for mobile use, such as paper claiming functions and determining what information to display for each paper.

7. Conclusion

This paper briefly discussed the status of institutional repository development in domestic university libraries, 详细介绍了 Tongji University's institutional repository service practices from perspectives of user needs analysis, characteristic service models, and service effectiveness; summarized three work experiences based on over four years of service, including data quality, usage details, and promotion methods; and finally explored several possible future service directions for institutional repositories.

University libraries are facing transformation, and the gradual rise of institutional repositories represents both challenge and opportunity. Leveraging institutional repositories can reshape the library's position within the university, helping it escape the awkward situation of marginalization. The basic attribute of institutional repositories is resource preservation. If only this function is considered, the repository operator is replaceable. Libraries should fully leverage their advantages, such as cataloging skills and subject services, utilize institutional repository data, integrate various information services, and provide value-added services to different user groups in a user needs-oriented manner. Only by strengthening service awareness and providing value-added services to various user groups can libraries make institutional repositories trusted and utilized across multiple scenarios, thereby playing a more important role in universities.

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

Source: ChinaXiv — Machine translation. Verify with original.