Current Status of Socialized Application Behavior of Institutional Repositories in Japanese Universities Postprint

Authors: Li Xi

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

Abstract

[Purpose / Significance] By collecting information on the social application behaviors of institutional repositories in Japanese universities, this study analyzes the current status of such behaviors. [Method / Process] Through literature review, we investigate the current research status on social application behaviors of institutional repositories, and by visiting websites of different institutional repositories, university intranet databases, and the National Institute of Informatics of Japan, we examine the relationships between institutional repositories and databases in other fields. [Results / Conclusion] Japanese university institutional repositories enhance their degree of social application through measures such as clarifying their positioning in policy documents, releasing specialized descriptive metadata formats for IR resources, applying OAI-PMH, and enabling data interoperability between databases. In the future, the objects of data interoperability will be extended to related fields such as the publishing industry to develop more in-depth data services.

Full Text

Research on the Socialized Application Behavior of Institutional Repositories in Japanese Universities

Li Xi
School of Computer Science and Information Technology, Northeast Normal University, Changchun 130117

Abstract

[Purpose/Significance] This study analyzes the current state of socialized application behaviors of institutional repositories (IRs) in Japanese universities

by collecting relevant information about their outreach and integration efforts. [**Method/Process**] Through literature review and systematic investigation of various IR websites, university library databases, and the National Institute of Informatics (NII) portals, this research examines the relationships between IRs and other domain databases. [**Result/Conclusion**] Japanese university IRs enhance their socialized application through several key measures: establishing clear positioning in policy documents, developing specialized metadata formats for IR resources, implementing OAI-PMH protocols, and enabling data interoperability across databases. Future developments will extend interoperability to the publishing industry and other related fields to provide more comprehensive data services.

Keywords: institutional repository; socialized application; open access; socialization
Classification Number: G250
Citation Format: Li X. Research on the socialized application behavior of institutional repositories in Japanese universities [J/OL]. Knowledge Management Forum, 2016, 1(6): 423-432. http://www.kmf.ac.cn/p/1/71/.

Institutional repositories (IRs) have emerged as a crucial implementation pathway for open access and are gaining increasing attention. Universities, as vital research and education institutions with diverse and numerous profiles, have made their IRs essential channels for achieving open access to resources. As universities deepen their understanding of open access and their demand for sharing academic resources grows, more institutions are attempting to build their own IRs.

However, since their inception just over a decade ago, IRs have generally lacked the visibility of traditional academic journals. Coupled with the complex intellectual property issues involved, IRs face significant challenges in resource acquisition. This resource scarcity, in turn, hinders efforts to enhance IR visibility, creating a vicious cycle that prevents IRs from fully realizing their potential. While intellectual property issues require long-term solutions through theoretical research and practical work based on existing legal frameworks, IRs can rapidly improve their visibility and academic standing through proactive measures such as opening their data, conducting promotional activities, and leveraging technical means to develop interoperability functions with other databases. These activities can be collectively termed as IR socialized application behaviors.

1.1 Basic Meaning

IR socialized application behavior refers to actions that enhance the openness of IRs and their data, improve social awareness, and establish connections with existing databases to actively meet information needs across various domains. The fundamental goal is to push resource information to users through multiple channels, facilitate barrier-free access to full-text resources, and enhance IR vis-

ibility. The data objects of IR socialized application behaviors include not only various types of data collected in IRs but also data not directly collected yet capable of establishing direct or indirect associations. Participants extend beyond IR managers to include IR content providers, various users, formal publishers, commercial database companies, and other social groups and institutions that can connect with IR data and promote its widespread application.

1.2 Related Requirements

IR socialized application behaviors are influenced by factors including institutional positioning, data structure and content, technical means, and promotional activities conducted by IR managers. At the macro level, IR socialized application requires IRs to integrate effectively into the broader social data environment, increase social participation, and leverage open access advantages to promote information resource sharing and exchange. At the micro level, effective implementation demands accurate role positioning and standardized data formats.

In terms of role positioning, IRs must clearly establish themselves as important pathways for open access, with highly open data being the key prerequisite. Open data encompasses not only content data but also metadata within IRs and usage statistics generated by various IR services. Generally, all data produced during IR operation and maintenance should be openly accessible, except for data involving personal privacy or resources under embargo periods, to meet diverse user information needs.

Standardized data formats mean that IR internal resource organization complies with international industry standards, employing universal metadata standards or adapted formats that accommodate institutional and IR resource characteristics while providing extended functionality. Standardized formats enable IRs to establish connections with other IRs within the same domain and facilitate interoperability with databases across domains, overcoming domain limitations. They also enable metadata harvesting through unified protocols, allowing regional aggregation of IR metadata and content data to support national-level IR services or search engine indexing, ultimately promoting IR development.

1.3 Distinction Between IR Socialized Application and Open Access

The socialized application model possesses characteristics that pure open access lacks. First, it emphasizes proactive openness and active service. By actively serving various types of IR users and potential users, IRs facilitate access to needed literature resources. Unlike library services that directly engage users, some IR socialized application behaviors—such as establishing data interoperability with other databases—do not directly involve IR users but still facilitate access and usage. Second, IR socialized application involves numerous participants, including all organizations or individuals who can facilitate IR data utilization, not just IR managers.

Therefore, IR open access does not equate to IR socialized application. The primary distinction lies in that the former remains at the stage of making various types of IR data openly accessible with limited application, while the latter attempts to break through institutional and domain limitations to establish connections with diverse datasets and develop data services. The socialized application model places greater emphasis on proactive service than open access, with open access serving merely as its foundation. Just as the contradiction between massive information environments and specific user needs prompted libraries to shift from passive to active service, IRs face similar challenges. To fully leverage their role and maximize the value of collected resources, IRs must move beyond the open access stage.

2.1 Domestic and International Research Status

Current research on IR socialized application primarily focuses on data interoperability, with data exchange protocols and system integration being common themes. Yao et al. discussed OAI-PMH functionality, examined existing problems, and proposed solutions []. Chen et al. explored the role of OAI-ORE in IR construction and function development based on OAI-PMH []. In Japan, Kojiro Shiroyuki and colleagues examined OAI-PMH concepts and detailed metadata exchange between Chiba University's academic repository and NII's metadata database []. Hayashi Yutaka discussed OAI-ORE as a successor protocol to OAI-PMH []. Wang et al. explored data interaction models between ORCID, iAuthor, and IRs, proposing a technical framework based on iAuthor and implementing it on Lanzhou University's IR platform [].

In system integration practice, UK university libraries, represented by University College London, have implemented the RIOJA project (Repository Interface for Overlaid Journal Archives), which developed tools integrating arXiv's self-archiving platform with the OJS (Open Journal Systems) electronic publishing system to improve resource collection and sharing efficiency []. In data interoperability, Yamamoto Tetsuya from Nagoya University conducted research on data conversion between different systems [], while Tanahashi Koreyuki and colleagues developed a name authority control system for inter-system connectivity [].

2.2 Research Methods and Data Sources

This study primarily employs literature review methods. When examining the status of Japanese university IR socialized applications, Office Visio software was used to create data utilization structure diagrams for visual representation. Information about Japanese university IRs was collected through the NII's Academic Institutional Repository Construction Support Project website, which provided details on IR construction and related research activities. Data on OAI-PMH implementation was gathered from the official OAI-PMH website and NII's "Metadata-Database Co-construction Project" portal. Practical

cases of IR system integration were collected from researchmap, Japan's largest researcher information platform, and various university IRs.

3.1 Japanese IR Development Environment

In the early 21st century, NII and other institutions launched the "Cyber Science Infrastructure" project to build a shared information resource network via the internet. To advance this initiative, NII began close collaboration with domestic universities, with research output dissemination network construction being a key component. University IRs were expected to play crucial roles in this process, leading to the inclusion of IR commissioning projects in the system development. Since 2005, NII has collaborated with universities interested in building IRs through commissioning projects, providing financial support for approved initiatives.

In 2013, the Japan Science and Technology Agency (JST) released its formal open access policy [], with IR utilization strategies included in the specific measures section. This policy accelerated open access adoption nationwide. J-STAGE (Japan Science and Technology Information Aggregator, Electronic), JST's digital magazine open access system, also clarified its positioning in its usage regulations: promoting the circulation and dissemination of digital scientific and technical literature domestically and internationally while developing an evolving information distribution infrastructure platform []. These policy documents naturally exerted positive influence on IR development as a key open access pathway.

Simultaneously, numerous university libraries expanded cooperation and exchange in digital resource co-construction and sharing. In October 2010, NII and the National, Public, and Private University Library Cooperation Committee signed a cooperation agreement on digital resource construction and utilization. Based on this, the Japan Alliance of University Library Consortia for E-Resources (JUSTICE) was established in April 2011 to ensure stable, continuous access to electronic resources needed for teaching and research. JUSTICE negotiates purchase and usage terms with publishers and copyright holders while providing support for electronic resource management systems to member libraries. It also organizes training sessions and publishes materials for librarians responsible for digital resources [], influencing university library digital resource construction from multiple perspectives including resources, platforms, and personnel. As an important component of digital resource construction, IRs have consequently gained increasing academic attention and influence.

3.2 Current Status of IR Construction in Japanese Universities

According to the 2015 "Academic Information Infrastructure Status Survey" conducted by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), statistics on IR construction in Japanese universities are shown in

[]. Based on MEXT's university statistics, approximately 56.5% of Japanese universities have established IRs, including 98.8% of national universities, 69.8% of public universities, and 48.6% of private universities. Additionally, about 24% of universities express willingness to build IRs. This demonstrates that Japanese university IRs have achieved considerable scale, yielded rich research and practical results, and operate in a relatively mature development environment. Studying their socialized application status can provide valuable references for IR theory development and reveal both advantages and shortcomings in practical implementation.

3.3 Research Projects Related to IR Socialized Application Behavior in Japanese Universities

Investigation reveals that between 2006 and 2011, more than two research projects annually addressed IR socialized application behavior. Representative projects are listed in []. The AIRway project (Access path to Institutional Resources via link resolvers) aims to establish a portal for centrally presenting open access resources from IRs and other databases, enabling users to quickly locate and obtain accessible links to resources within authorized databases, thereby reducing acquisition time []. The experimental development project for constructing a cooperation model between repositories and electronic publishing, undertaken by Waseda University, Kyoto University, Hiroshima University, Nagasaki University, and Saga University, developed a Japanese-language digital publishing system with publication review functions to facilitate faster open access inclusion of bulletin papers and other internally published materials in IRs []

3.4 Structural Model of IR Socialized Application in Japanese Universities

IR socialized application behaviors are diverse, and any social group that can promote IR resource utilization and enhance IR visibility should be included in this model. Current participants in Japan's IR socialized application ecosystem include: various IRs and auxiliary databases such as the IRDB performance analysis system and national-level IR service JAIRO; researcher information platforms including the national researcher information database researchmap and university-level researcher databases; libraries such as the National Diet Library (NDL); academic information portals like J-GLOBAL; and search engines such as Google. The structure is illustrated in [Figure 1: see original paper], where solid arrows represent established, feasible data interoperability behaviors, and dashed arrows indicate behaviors still under exploration.

University IRs and other campus databases (e.g., researcher information databases) form the foundation of IR socialized application activities. IRs can directly serve users or provide access, retrieval, and resource acquisition through IRDB or JAIRO. They can also establish data interoperability with

other campus databases to offer diverse access methods: users can access IRs through specific links when visiting other campus databases or via information platform portals like researchmap and J-GLOBAL that are based on campus researcher and research output databases. Through OAI-PMH, search engines can harvest IR metadata, enabling users to retrieve IR resources directly via search engines and access them through provided links.

Following policy changes in doctoral thesis submission in 2013, an increasing number of IRs began collecting their institutions' doctoral theses. Previously, NDL primarily handled doctoral thesis deposit work. Through data sharing between NDL and IRs, NDL automatically collects metadata for theses deposited in IRs, allowing users to search IR doctoral thesis resources through NDL's OPAC function.

Collaboration with publishing domain databases represents a crucial aspect of advancing IR socialized application. Research in this area began early, with universities such as Waseda and Nagoya conducting relevant studies. Since 2006, Waseda University has led the "Electronic Publishing (Editorial Review) System Development" project under NII's IR commissioning program [], aiming to streamline the registration process for bulletins and internal publications in IRs and enable their electronic publication. Between 2006 and 2009, the project progressed continuously, establishing a dedicated website, and by 2008, a publishing system using OJS was completed. During electronic resource publication, the system provides submission functions for creators and allows librarians to extract data and register publication information in IRs upon completion. This data interaction between IRs and publishing systems enables eligible outputs to be included in IRs at the time of formal publication, advancing open access to the moment of creation and even extending it to pre-publication works, thereby prolonging the open access timeframe.

The ERDB-JP database, constructed collaboratively by numerous academic societies, publishers, and universities, serves as an open knowledge space centralizing data on electronic journals and resources published in Japan. Originating from the Electronic Resources Database (ERDB) project launched in 2012 and jointly promoted by NII and multiple university libraries, ERDB-JP plans to automatically harvest IR data through programmed crawlers, recording URLs in its user information database. This initiative, launched in June 2016, enables efficient and widespread dissemination of data related to bulletins and internal publications collected in IRs, expanding utilization pathways and creating favorable conditions for increasing the use of IR content []. Additionally, researchmap and JAIRO Cloud are actively cooperating, with the data interoperability development between Shinshu University's researcher directory database and its IR serving as an exploratory practice based on this collaboration. Such interactions with systems like researchmap explore methods for IR cooperation with other social information systems, accumulate collaborative experience, and positively influence IR integration into the broader social data environment.

Diversified access pathways increase the likelihood of user visits, reduce time

costs for resource acquisition, and facilitate more convenient IR utilization, which is extremely beneficial for IR development. However, despite overall rapid development momentum, certain areas still exhibit shortcomings with substantial room for improvement. While the vast majority of Japanese university IRs have achieved open access to resources with minimal barriers, some IRs fail to develop further services after resource collection and open access implementation, only functioning when users visit IR websites to browse and download resources. Practices in data interaction between IRs and the publishing domain remain very limited, with related progress still at the experimental exploration stage.

4. Measures Related to Socialized Application Behavior

As complex information systems, IRs require multifactorial support to realize socialized application behaviors. Four particularly important factors are: IR application policy (i.e., IR self-positioning), IR data format, metadata harvesting protocol application, and IR interoperability with related systems. The first three constitute the foundation of the socialized application model, while interoperability represents its concrete manifestation. These four factors are currently relatively implementable and thus form the focus of this analysis.

4.1 IR Self-Positioning

By investigating Japanese university IR websites and libraries, this study collected 291 operational policy documents and IR positioning descriptions. These documents typically define institutional IR positioning prominently, often in the first article. For example, Hokkaido University's IR operational policy document "Regulations on the Collection and Management of Hokkaido University Academic Achievements" states in Article 1: "to be publicly available both on and off campus through the internet," with detailed provisions in supplementary regulations []. The collected positioning information includes both simple introductions containing terms like "public," "free," and "gratuitous," and fair use guidelines formulated by IR managers based on copyright law that permit usage within specified scopes.

4.2 IR Data Format

For Japanese university IRs to have their metadata harvested by the IRDB system, two conditions must be met []: first, a standardized URI for accurate IRDB location; second, compliance with specific metadata standards—namely, support for the OAI-PMH protocol and the Junii2 standard.

Junii2 is a metadata format released by NII in 2006, with Version 1.0 published in March 2010 and Version 3.1 revised in March 2013. JAIRO, JAIRO Cloud, and the IRDB performance analysis system all rely on Junii2 for metadata harvesting, integration, aggregation, and service provision. Junii2 comprises 64 descriptive elements, each with eight recording items including title, necessity,

repeatability, maximum/minimum occurrence, and input format []. The format facilitates interoperability with numerous international classification systems, as evidenced by dedicated sub-elements for major classification schemes and thesauri. For instance, the "Subject" element contains 11 sub-elements for systems including the National Diet Library Classification, NDC, BSH, LCC, and DDC. The "ID" element includes sub-elements for PubMed ID, DOI, NII Article ID, and Ichushi ID. Such designations enable resource description according to international standards. Junii2 also provides English content for each element, establishing connections between native language vocabularies and foreign metadata formats to lay the linguistic foundation for metadata aggregation and system interoperability, while facilitating XML-based publication of resource metadata.

4.3 Application of Metadata Harvesting Protocol OAI-PMH

Support for OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) is essential for IRDB metadata harvesting. Developed and maintained by OAI, the protocol's latest version is 2.0, released in 2002. Based on HTTP, OAI-PMH can transmit various XML-formatted datasets, enabling harvesting of individual metadata records, frameworks, or entire datasets from target IRs as needed. Without OAI-PMH, automated metadata harvesting and large-scale open application would be impossible. NII launched its metadata-repository co-construction project in October 2002 to explore practical OAI-PMH applications, with Chiba University conducting simultaneous trials [].

However, OAI-PMH has limitations, addressing only metadata harvesting and transmission without standardizing content data, which restricts its application scope—a constraint that becomes increasingly apparent as web resource types grow more complex. OAI-ORE (Open Archives Initiative Object Reuse and Exchange) was developed to address these shortcomings, enabling digital resource transmission between platforms beyond metadata. OAI-ORE application in IRs remains at the theoretical research stage.

4.4 Interoperability Between IRs and Related Systems

Interoperability with related systems represents a concrete manifestation of IR socialization, enhancing resource utilization. Interoperable systems include various research output databases, researcher information databases, publishing domain databases, and search engines both on and off campus.

Two exemplary cases are Kanazawa University's education and research performance database [] and Shinshu University's researcher directory [].

Kanazawa University's database features an interoperability module enabling researchers to send research outputs directly to the IR from their personal accounts without re-entering information. After simple metadata description, information can be transmitted to the IR, where managers can review submissions upon login, checking for copyright issues before confirming inclusion. Managers may also modify description information before finalizing records.

In Shinshu University's researcher directory, a researcher's profile displays organizational information, position, research fields, projects, affiliations, and outputs. If an output is included in Shinshu University's IR, a full-text link appears in the corresponding output record, redirecting to the IR page. Conversely, IR output pages contain reverse links to the researcher directory, enabling navigation to comprehensive researcher profiles. This interoperability reduces researchers' input workload during submission, simplifies management processes, and creates complementary advantages: IRs provide full-text resources while researcher databases offer comprehensive personnel information, indirectly enabling literature provision capabilities for the latter and researcher identification capabilities for the former.

5.1 Broader Scope of Socialized Application Activities

Currently, IR socialized application extends beyond IR domain databases to include researcher information databases and research output databases. However, substantial potential exists for broader data interoperability relationships with publishing domains, commercial databases, and academic information portals. In publishing, academic outputs generate substantial publication data during formal publication. Commercial databases contain descriptive data for officially published versions of academic works. Academic information portals, as comprehensive platforms, publish diverse information about academic outputs. Establishing associations with publication data can facilitate IR managers' copyright and publication status reviews. Connections with commercial databases can help IR users access officially published content, particularly valuable for research data and pre-peer-review manuscripts in IRs. Linkage with academic portals can enrich information types, improve overall service capabilities, and provide more convenient access channels for IR resources, thereby enhancing IR visibility.

5.2 More In-Depth Socialized Application Services

Although IRs have achieved interoperability with researcher information and research output databases, the scope remains limited by copyright restrictions, preventing comprehensive information sharing. As open access advances, more resources will become available through various fair use forms, with full texts potentially becoming open access. Services enabling IR users to access data from other interoperable databases will deepen. With IRs collecting increasingly more outputs, discovering relationships among them becomes more important. JAIRO already enables searching over 2 million output records, providing a substantial data foundation for knowledge and relationship discovery. Combined with developing information networks between IRs and researcher/output databases, this enables knowledge management activities at national, institutional, and individual levels. Researchmap already offers personal knowledge management capabilities and can even function at the national level in certain studies. Having established connections with university-level researcher

databases, researchmap has achieved indirect data interaction with IRs. As technology matures, the information network between IRs and researchmap will become more complete, enabling more comprehensive and in-depth knowledge management activities at institutional or national levels.

References

[] Yao X, Zhu Z, Lu L, et al. Research on OAI interoperability data synchronization strategies for institutional repositories [J]. Modern Library and Information Technology, 2014(3): 14-

[] Chen X, Zhang Z, Bai H. Research and implementation of OAI-ORE in institutional repositories [J]. Modern Library and Information Technology, 2010(11): 69-

[] Shiroyuki K, Sugita S, Asajima H, et al. Attempts to construct academic institutional repositories in Japan: Focusing on cases of Chiba University and National Institute of Informatics [J]. Journal of Information Science and Technology, 2004, 54(9): 475-482.

[] Hayashi Y. ResourceSync: Successor specification to OAI-PMH [EB/OL]. [2016-09-10]. http://current.ndl.go.jp/ca1845.

[] Wang H, Huang J, Liu X. Research on ORCID application in institutional repository construction [J]. Library and Information Service, 2015, 59(17): 134-140,133.

[] University College London. RIOJA [EB/OL]. [2016-09-10]. http://www.ucl.ac.uk/ls/rioja/.

[] Yamamoto T. Developing a universal data conversion framework [J]. Annual Report of Nagoya University Library, 2007(6): 71-

[] Tanahashi K, Yamamoto T, Kajita T, et al. Development of an author name authority directory system for inter-system connectivity [J]. Annual Report of Nagoya University Library, 2007(6): 63-

[] National Institute of Informatics. Next-generation academic content infrastructure co-construction project commissioning guidelines [EB/OL].[2016-09-10]. https://www.nii.ac.jp/irp/rfp/2006/pdf/rfp.pdf.

[] Japan Science and Technology Agency. JST policy on open access [EB/OL]. [2016-09-10]. http://www.jst.go.jp/pr/intro/pdf/policy_{openaccess}.pdf.

[] Japan Science and Technology Agency. J-STAGE terms of use [EB/OL]. [2016-09-10]. https://www.jstage.jst.go.jp/pub/html/AY04S090_{ja}.html.

[] National Institute of Informatics. Overview of Japan Alliance of University Library Consortia for E-Resources [EB/OL]. [2016-09-10]. http://www.nii.ac.jp/content/justice/overview/.

[] Ministry of Education, Culture, Sports, Science and Technology. Academic information infrastructure status survey 2015: University library section on institutional repositories [EB/OL].[2016-09-10].http://www.e-stat.go.jp/SG1/estat/List.do?bid=000001070838&cycode=0.

[] National Institute of Informatics. Academic institutional repository construction support project - commissioned projects [EB/OL]. [2016-09-10]. https://www.nii.ac.jp/irp/rfp/.

[] Hokkaido University Library. AIRway project: Access to institutional re-

sources via link resolvers [EB/OL]. [2016-09-10]. http://airway.lib.hokudai.ac.jp/index_{ja}.html.
[] Waseda University. Experimental development for establishing a cooperation model between repositories and electronic publishing [EB/OL]. [2016-09-10]. http://www.wul.waseda.ac.jp/ir/epubs/.
[] Kataoka M. Past efforts and future plans: Automatic data integration with JAIRO Cloud and system updates [EB/OL]. [2016-09-10]. http://www.nii.ac.jp/csi/openforum2016/track/pdf/20160526_{ERDBJP}4_{kataoka}.
[] Hokkaido University Library Committee. Regulations on collection and management of Hokkaido University academic achievements [EB/OL]. [2016-09-10]. http://eprints.lib.hokudai.ac.jp/siryo/unyo.pdf.
[] National Institute of Informatics. Conditions for IRDB harvest target IRs [EB/OL]. [2016-09-10]. http://www.nii.ac.jp/irp/archive/system/irdb_{harvest}.html.
[] National Institute of Informatics. Metadata format junii2 (version 3.1): Input content list for each data element [EB/OL]. [2016-09-10]. https://www.nii.ac.jp/irp/archive/system/pdf/junii2_{{{elements}}*{{guide}}}*{ver3}.1.pdf.
[] Kanazawa University. Kanazawa University researcher information [EB/OL]. [2016-09-10]. http://ridb.kanazawa-u.ac.jp/public/.
[] Kanazawa University. Kanazawa University academic information repository [EB/OL]. [2016-09-10]. http://dspace.lib.kanazawa-u.ac.jp/dspace/.
[] Shinshu University. Researcher directory [EB/OL]. [2016-09-10]. http://soar-rd.shinshu-u.ac.jp/search/index.html?lang=ja.
[] Shinshu University. Institutional repository [EB/OL]. [2016-09-10]. https://soar-ir.repo.nii.ac.jp/?lang=japanese.

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