Creating an Article Repository using LibSys: A Case Study of JRD Tata Memorial Library

K.S. Chudamani* and H.C. Nagarathna

JRDTML, Indian Institute of Science, Bangaluru-560 012
*E-mail: ksc@library.iisc.ernet.in

ABSTRACT

Article Repositories are an integral part of present day digital libraries permitting global access to the scholarly publications. An article repository system can be created using LibSys software, and help LibSys users collaboratively create a federated search engine to share resources among themselves. This article highlights the system used by the JRD Tata Memorial library of Indian Institute of Science, Bengaluru for providing open access to articles published by its faculty.

Keywords: IISc, institutional repository, open access, LibSys, knowledge management, web OPAC, resource sharing

1. INTRODUCTION

Institutional article repository is an integral part of present day digital libraries. They provide a simple web-based mechanism to researchers to deposit and access their research publications. Institutional repository (IR) capture articles, books, reports, etc. with a view to considering the long term preservation of an institutions scholarly output. Articles from Indian Institute of Science (IISc), Bengaluru, feature significant textual content and annotated metadata like resource, date of publication, author, title, year issue no., vol no., etc. In addition, keywords, subject categories, etc., are available through both online (web based) and offline clients. Creation of such an article repository is possible by using LibSys software, and involves a number of knowledge management (KM) issues.

2. KNOWLEDGE AND KNOWLEDGE MANAGEMENT

Knowledge and knowledge management (KM) have very close relation in many ways. According to Davenport and Prusak, "Knowledge is a mix of framed experiences, values, contextual information, expert might, and grounded intuition that provides an environment and framework for evaluating incorporating new experiences and information…. It often becomes embedded not only in documents or repositories but also in organisational routines, processes, practices, and norms……" [1].

Production of new explicit knowledge is dependant on the results of several business transformations where intelligent data and text mining algorithms, such as decision trees, neural networks, case-based reasoning, genetic algorithms, etc., applies to structure or unstructured data and information. Data/text mining is concerned with detecting new information in large databases/large collections of text documents. Typical data/text analysis includes identifying trends, discovering relationships among terms and documents, categorizing or condensing information.

In simple terms, KM is management of knowledge. According to Tiwana [2], it enables:

❖ Competitive advantage.

❖ Business goals.

❖ Business competencies.

❖ Innovative products and or services.

❖ Enhances relationships.

❖ Improves work practices and processes.

All these tasks mainly need to handle tacit knowledge (TK) and explicit knowledge (EK). Usually, TK is not easily available for dissemination. It is in user's perception. In addition, TK is moving and enables competition. On the other hand, Information Retrieval

Systems (IRS) enables a simple, web-based mechanism to researchers to deposit (self-archive) and accesses their research publications. They can also incorporate TK to some extent by including documents like technical reports, patents, etc. There are many world-renowned open source IR softwares such as Eprints, DSpace, FEDORA, ARNO, i-TOR, CDSWare, etc., available free of cost, and have a host of features, unique facilities and excellent capabilities, which the users could explore and experiment upon. The need for deploying interoperable open access IRs is essential for knowledge transmission. Knowledge requires sharing and IT supports sharing. IT has made information sharing possible through article repositories, which are OAI-PMH compliant or otherwise. LibSys is an OAI-PMH compliant as it is MARC compatible. Further, Tiwana points out that KM is not a technology problem, but is a process problem. This implies that expert knowledge differs from organization to organization. Libraries tend to tag knowledge sources using their attributes. Some such attributes for entering data into databases using libsys are:

- ❖ Bk: books

- ❖ Cp: conference proceedings

- ❖ Th: theses and dissertations

    The above three can be merged and searched.

Knowledge management of institutional publications has gained ground in these days of open access movement. As pointed out by Sreekumar [3], a digital repository is created to provide access to institutional faculty lectures, class notes, publications, etc. Sreekumar [3] and Poornima [4] have described IR developed in their institutions using Dspace and Eprints, respectively. For articles, separate database, which is completely indexed, modified, and searchable can be created using Libsys. There are many more attributes, based on which one can build the collection. The method of sharing the resources of various organizations through the LibSys software is possible. This can enable Libsys users to create collaboratively a federated search engine.

Resource sharing as already pointed out is an important activity carried out by librarians. So far, they were usually restricted to inter lending of books, journals, document delivery service, i.e. supply of articles and chapters from books to users on minimum payment basis. The advent of repositories has enabled databases of articles, books, etc., published by the faculty, staff, and students searchable, browseable, and downloadable. This is a first step towards free exchange of bibliographic information. A home page for searching available repositories can be created using Google or Scopus search engine as displayed in National Centre for Science Information (NCSI) and IISc library home page. A federated search is in use for information mining.

However, in this article we are looking at creating article repository only for providing open access for articles published by IISc faculty. The article describes the creation of an article repository using "article indexing" module in LibSys. Before describing the actual repository creation methodology, it is worthwhile to look at some of the existing article repositories in India.

3. INSTITUTIONAL REPOSITORIES IN INDIA

According to Gayatri Doctor and Smitha Ramachandran [5], there are 16 digital repositories in India. Out of these, 10 are based on Dspace and six on GNU Eprints

According to Prasad [6], about 30 institutions in India are using DSpace (3rd position in the world). Some of the important features of Dspace repository are:

(i) Dublin core metadata.

(ii) Open archives initiative compliance.

(iii) Web interface which is customer enabled.

(iv) All types of documents such as books, theses, articles, etc. are accepted.

(v) simple workflow processes.

(vi) Full text searching.

JRD Tata Memorial Library (JRDTML) of IISc, Bengaluru is one of the institutional libraries, which is also using these features in its IR.

4. JRDTML REPOSITORY

The library of IISc is one of the best scientific and technical libraries in India. Apart from the main library it has departmental libraries also. In 1995 the library was renamed as JRDTML. IISc researchers now have online access to electronic journals and bibliographic citation databases under the INDEST-AICTE Consortium, and have access to fairly a large number of journal archival content from the resources.

JRDTML has been using the LibSys with all the modules for the library house keeping operations. The library's online catalogue database has more than 1.50 lakh records of books (bibliographic information) and information about 1, 72, 000 back volumes of periodicals. Using LibSys OPAC, users can search the Library Online Catalogue by author, title, subject, and keywords. The users can know the latest additions of periodicals and books through current awareness service. List of all the articles indexes during a week, i.e., articles published by the institute's faculty and students (collected from various databases) is generated. The JRDTML repository uses a database search technology called article indexing in

LibSys. It focuses on efficient storage and fast retrieval of short text entities. Basic functionality of LibSys has an extended clustering mechanism that applies to search results and a similarity search mechanism based on the articles by IISc Authors. There is also a separate repository developed by National Centre for Science Information for the users.

4.1 Eprints Archiving Facility (eprints@iisc)

NCSI has set up an Eprints archiving facility for the IISc research community, using the GNU Eprints software developed at the University of Southampton. This service enables the Institute to archive their pre-prints, post-prints and other scholarly publications.

In keeping with the objectives of the Eprints movement and the Free Online Scholarship (FOS) movement, we expect this service to facilitate the Institute researchers in self-archiving and long-term preservation of their scholarly publications, provide easy access to these publications worldwide and improve impact of their research. While eprints@iisc is accessible by anybody, submission of documents to this archive is limited to the IISc research community only. This repository and the library repository described are complementary to one another in the sense that the article repository of the library lists forthcoming articles and the NCSI repository collects cited articles. The repository creation process using Dspace, Eprints, etc., needs acquisition of knowledge of installing these utilities using tomcat, apache web server, etc. It also requires the knowledge of using the software meticulously by giving passwords to users. Uploading the article from a remote point is possible. Interoperability among LibSys ensures that the fields are standardised within LibSys framework. Libsys framework can crosswalk with MARC21 format. This can be used to even store other types of knowledge sources like lecture notes, etc.

5. AIMS OF THE IISC LIBRARY REPOSITORIES

Aims of the IISc library repository are to:

❖ Improve interoperability and cooperation between IISc author's article, and index the running digital repositories.

❖ Provide a tool that could be used for the repository operation to raise the visibility and importance of the LibSys repository with in the campus.

❖ Provide a detailed description of the institution author article requirements of certified document and publication services.

❖ Document publication services as an institutional repository.

❖ Enable development services and information exchange.

❖ Improve interoperability and cooperation between IISc author articles index by running cross-indexing digital repositories.

❖ Provide a tool that in use by the repository operation to raise the visibilities, important of the LIBSYS repository.

Article repository proposes to facilitate improved access, visibility and impact of IISc research output through the establishment of a network. It is an open access to digital research articles. It provides technical support to R&D organizations and universities in the country to set up their institutional research repositories.

6. LIBSYS REPOSITORY OF IISC

6.1 Creation of the Repository

According to Gayathri Doctor and Smitha Ramachandran [5], factors for consideration while developing an IR are: content, technology, users/uses, legal aspects, sustainability, and funding. As already pointed out, the content for IISc article repository are the articles published by its faculty, the strategies adopted were different from those of the others; usually, when creating an article repository, articles of the authors are searched from various sources and are made available through the article indexing part of the software. ICT technology enables the data to be available through the web so that users accessing the library web-OPAC anywhere can access bibliographic description along with the content. Since this is an institutional repository, it does not involve the copyright problems. The steps involved in this are:

❖ Duplicate checking.

❖ Data entry.

❖ Keyword assignment.

❖ Editing.

❖ Developing weekly list of additions.

The sustainability of the project is dependant on the willingness of the staff or making the staff available for the project.

6.2 Searching, Indexing and Retrieval

Any information retrieval system such as the OPAC designed for any automated library has characteristics for searching by author, title, content, keyword, etc. The OPAC for any sources should enable browsing, searching, multimedia linking to e-books, e-journals, etc. If the user types in a keyword "structure" what exactly

user means or in what context it is being searched is not known. The user proceeds with the search process in a focused manner based on the display. If it fails to produce satisfactory results, user makes use of other search strategies. User may use hierarchical searching, content searching or combine all these search strategies. Most of these searching capabilities are established techniques in library and information science field. Therefore, this presentation confines to only multimedia aspects.

6.3 Multimedia

Multimedia, as found in fancy websites, is not the real captivating feature. It has an important function, i.e., relating databases and websites. Multimedia relates to video and text that surpasses all limitations of language.The article indexing system of IISc provides the facility to maintain a separate database containing the articles in journals received in the library. It is also available on OPAC. LibSys focuses on giving guidance to IR operating to help them to support the open access concept and supplying pre-and post-prints of journal articles index that have already been published in scientific journals. A maintenance functions provides for building up the articles database in the LibSys software: for addition of a new article, modify details of an article, removal of an article, list article details, and option to print the same.

6.4 Advanced Search Options and Data Mining

Database searches for the article in the database is searchable by browsing any of the following: author, title, sub/keyword, searching by title keywords, and combination searches using Boolean operators (based on words from any field). The keyword forms the basis of

ontology. Classification codes are not a search feature as the repository has placed emphasis on keywords. Figure 1 shows the screen shot of web-OPAC window for article index.

The bibliographic data is present in various databases, which the institute library subscribes and hence there is no need for any approval by the author. When a user detects an article of interests in the article database within the web-OPAC database options, he/she can open the article in a separate window. Article display shows metadata directly with in the article text. For example, after a recognised author's name, a small "e-resource" is displayable and clickable to display the article for viewing.

6.5 Article Visualisation

Article visualisation increases the visibilities and ease of use of open access scientific journals, thereby promoting their increased usage and impact. Setting up open archives is becoming a part of library activities, as they are now more widely known as IR.

6.6 Latest Additions in the Article Database

The bibliographic function enables the generation of a list of articles on a specific subject. The selected entries are stored in a user defined 'Biblio' file.

7. ADVANTEGES

The following are the advantages of the IR:

❖ Helps organisation's in building intellectual capital.

❖ Enables a single access point for the research output of an institution.

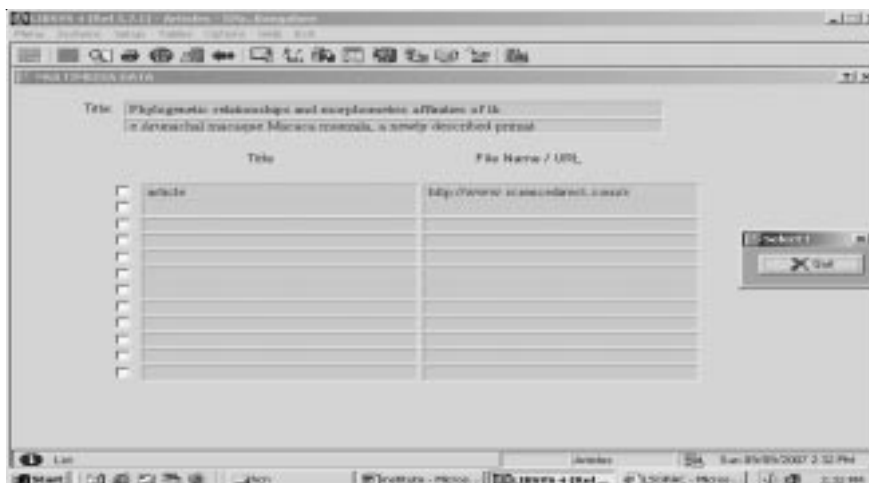Figure 1. Screen shot of LibSys article indexing page of IISc.

Figure 2. Screen shot of multimedia screen.

Figure 3. Screen shot of web OPAC.

❖ Shows actual direction in its research activities.

❖ Saves the time of instant generation of research report.

❖ Enables long term preservation of its scholarly materials.

8. CONCLUSION

Libraries who have procured Libsys software can easily create the IR without the hassles of implementing repository software. Since there are more than 1000 LibSys installations in the country, they can easily use the article indexing facility in LibSys and create repository. In addition, they can generate weekly list of additions, compile a bibliography on a subject, etc.

REFERENCES

1. Davenport, T.H. & Prusak, L. Information ecology, mastering the information and knowledge management, NY, OUP, 1997. pp. 3.

2. Tiwana, Amrit. Knowledge management toolkit, Pearson Education Asia, New Delhi, 2000. Chp 6.

3. SreeKumar, M.G., *et al.* Institutional repositories for knowledge management in academic and research Institutions. ICSD 2007, Bengaluru, pp. 260-273.

4. Poornima Narayan, *et al*. Instiutional repository at NAL: A case study, ICSD 2007, Bengaluru, 2007, pp. 274-84.

5. Doctor, Gayatri & Ramachandran, Smitha. Implementing an institutional repository for knowledge sharing in a management institution. *In* International conference on future of networked environment (IKONE), 3-5 September 2007, Bangalore (Poster Session).

6. Prasad, A.R.D. Institutional repositories: An over view (a ppt presentation). *In* Open Access and Institutional Repositories. One-day Workshop-cum-Seminar, INSA, 13 May 2004, New Delhi.