14. Collection Development in Digital Information Repositories in India
Anup Kumar Das1 and B K Sen2 and Chaitali Dutta3

ABSTRACT
The institutional repository (IR) is a contemporary concept that captures and makes available through Internet and intranet the institutional research output and other relevant documents to the users by way of digitizing the output. The IRs have already started emerging in India. This study highlights the importance of IR, delineates the scope and methodology, projects the findings. Most of the repositories are using open source information repository software like DSpace, Greenstone Digital Library Software and GNU EPrints. It is observed that generally documents like theses and dissertations, seminar papers, journal articles, etc., are being found more in the repositories. Some of the problems of the repositories have been highlighted and suggestions offered.

KEYWORDS
Institutional repository; Information repository software; Digital libraries; Collection development; India.

INTRODUCTION
In India there are a number of internationally reputed institutions, which are producing a good number of research documents that are expanding the frontier of knowledge and scope of technological innovation. The recent phenomenon of outsourcing of scientific research, applications development as well as business-intelligence-related research to India, has generated a good deal of enthusiasm in advanced studies. Recently some internationally famous journals published cover stories on Indian scientific and technological research. The success and achievements of these institutions can give rise to promising research environment in India that may attract sponsored or collaborative research in all major disciplines of study. These institutions essentially use modern information and communication technologies for information management and dissemination. Other than their research activities, these institutions also produce intellectually mature graduates and scholars in some scientific, technological and business disciplines. Other areas of studies in India are also getting international attention proportionately. Some of these institutions provide access to their research documents and learning materials initially to the Indian scholars in other institutions as well as to external scholars in institutions across the globe. The sharing of knowledge may lead to further development in the same discipline or related disciplines. Institutional repository, which may be called an extension of digital library, is now becoming a platform for the sharing of knowledge.

INSTITUTIONAL REPOSITORY
The Institutional repository (IR) is a contemporary concept that captures and makes available as much of institutional research output as possible to the users. It is a sort of a database of digital information resources, accessible through Internet or Intranet. In the first instance this might include electronic versions of documents such as research papers, project reports, patents, theses and dissertations [1]. It may also include many of the digital assets generated by an institution such as working papers, lectures, conference proceedings, learning objects, administrative documents, course notes, etc. The learning objects may include among others study materials, assignments, question papers, audio-video materials and multimedia presentations such as interactive e-learning modules. Advancement of information and communication technologies has made the institutional repository a reality. Institutions both in developed as well as developing countries have started establishing institutional repositories.

Importance of Institutional Repository
Institutional repository is the marquee of an institution to the world, where institution displays its worthwhile research programmes, projects, and initiatives to the broad spectrum of audience in the
world. An institution outreaches its findings that in turn encourage other institutions and organizations to collaborate and to share their knowledge, expertise and skills. An institutional repository offers seamless access to documents that reflect past and present research interests of the institution as well as its future research goals. It makes the publications more usable by contemporary and future scholars as well as other professionals like policy makers and social workers. The pace of scholarly communication would be highly accelerated if the IR holds research papers, research reports, etc. as soon they are made public. This also helps publications in receiving more citations, since the research findings are quickly available to the fellow scholars. The institutional repository can be used throughout the institution and collaborative institutions.

Some institutional repositories in India are only providing access to metadata to the external communities who are accessing these repositories through Internet, whereas internal members who are accessing these repositories through Intranet are getting access to full-text information besides metadata. These restrictions exist due to various reasons involving copyright issues, bandwidth issues, permissions from the grant providing agencies (GPAs), and so on. In India, there are a number of research grant providing agencies. Sometimes some areas may overlap with each other. Using institutional repositories, the GPAs can evaluate the novelty of a research proposal and come to know whether any study has been already undertaken in a particular area or discipline.

**SCOPE AND METHODOLOGY**

The present study covers selected institutional repositories in India accessible through Internet. Most of the repositories are using open source information repository software like DSpace, Greenstone Digital Library Software and GNU EPrints. Some institutions presently are not using any IR software, but providing direct access to their documents through their respective websites. For this study authors have visited respective institutional repositories using Internet and gathered information that are presented in Table 1. The Figures 1 to 4, depicted here, are drawn from the data as on 1st July 2005. (Footnotes)

1. Faculty of Library and Information Science, School of Social Sciences, Indira Gandhi National Open University, New Delhi - 110068, India
   Email: anupdas2072@hotmail.com

2. 80, Shivalik Apartments, Alaknanda, New Delhi - 110019, India
   Email: bksen@ndb.vsnl.net.in

3. Department of Library and Information Science, Jadavpur University, Kolkata - 700032, India

**Table 1: Selected Institutional Repository Initiatives in India**

<table>
<thead>
<tr>
<th>Name of IR</th>
<th>Institution</th>
<th>Website address</th>
<th>Type of Documents Available</th>
<th>Whether Full text</th>
<th>Number of documents as on 15/07/2005</th>
<th>Software Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDL: Librarians’ Digital Library</td>
<td>Documentation Research and Training Centre, Indian Statistical Institute</td>
<td><a href="https://drtc.isibang.ac.in">https://drtc.isibang.ac.in</a></td>
<td>Articles, conference papers, theses</td>
<td>Yes</td>
<td>188</td>
<td>DSpace</td>
</tr>
<tr>
<td>ICRIER</td>
<td>Indian council for Research on International Economic Relations</td>
<td><a href="http://www.icrier.org/publications.html">http://www.icrier.org/publications.html</a></td>
<td>Working papers, lectures</td>
<td>Yes</td>
<td>180</td>
<td>-</td>
</tr>
<tr>
<td>Repository Name</td>
<td>Institution</td>
<td>URL</td>
<td>Types of Content</td>
<td>Count</td>
<td>License</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------------------------------------------</td>
<td>----------------------------------------------------</td>
<td>-------------------------------------------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>eprints@iitmK</td>
<td>Indian Institute of Management, Kozhikode</td>
<td><a href="http://eprints.iitm.ac.in">http://eprints.iitm.ac.in</a></td>
<td>Articles, conference papers</td>
<td></td>
<td>GNU EPrints</td>
<td></td>
</tr>
<tr>
<td>ETD at IISc</td>
<td>Indian Institute of Science</td>
<td><a href="http://etd.ncsi.iisc.ernet.in/">http://etd.ncsi.iisc.ernet.in/</a></td>
<td>Theses and dissertations</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>ePrints</td>
<td>Indian Institute of Science</td>
<td><a href="http://eprints.iisc.ernet.in">http://eprints.iisc.ernet.in</a></td>
<td>Articles, conference papers</td>
<td></td>
<td>GNU EPrints</td>
<td></td>
</tr>
<tr>
<td>Prabhavi</td>
<td>Indian Institute of Science</td>
<td><a href="http://vidya-mapak.ncis.iisc.ernet.in/cgi-bin/library">http://vidya-mapak.ncis.iisc.ernet.in/cgi-bin/library</a></td>
<td>Articles, conference papers</td>
<td></td>
<td>Greenstone</td>
<td></td>
</tr>
<tr>
<td>ETD @IITB</td>
<td>Indian Institute of Technology, Bombay</td>
<td><a href="http://www.library.iitb.ac.in/~mnj/gdol/cgi-bin/library">http://www.library.iitb.ac.in/~mnj/gdol/cgi-bin/library</a></td>
<td>Theses and dissertations</td>
<td></td>
<td>Greenstone</td>
<td></td>
</tr>
<tr>
<td>Eprints@IIT Delhi</td>
<td>Indian Institute of Technology, Delhi</td>
<td><a href="http://eprint.iitd.ac.in/dspace">http://eprint.iitd.ac.in/dspace</a></td>
<td>Convocations addresses, faculty research publications</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>ETD @IIT Kanpur</td>
<td>Indian Institute of Technology, Kanpur</td>
<td><a href="http://library.iitk.ac.in:8080/examples/thesis/index.html">http://library.iitk.ac.in:8080/examples/thesis/index.html</a></td>
<td>Theses and dissertations</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>INSA Digital Library</td>
<td>Indian National Science Academy</td>
<td><a href="http://drtc.isibang.ac.in/insa/">http://drtc.isibang.ac.in/insa/</a></td>
<td>Articles, books, seminar reports</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>Digital repository of NCL</td>
<td>National Chemical Laboratory</td>
<td><a href="http://dspace.ncl.res.in">http://dspace.ncl.res.in</a></td>
<td>Theses, patents, project reports</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>OpenMed @NIC</td>
<td>National Informatics Centre</td>
<td><a href="http://openmed.nic.in">http://openmed.nic.in</a></td>
<td>Articles, conference papers</td>
<td></td>
<td>GNU EPrints</td>
<td></td>
</tr>
<tr>
<td>DSpace @nitr</td>
<td>National Institute of Technology, Rourkela</td>
<td><a href="http://dspace.nitrkl.ac.in/dspace/">http://dspace.nitrkl.ac.in/dspace/</a></td>
<td>Conference Papers, journal articles, preprints</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>Vidyanidhi</td>
<td>University of Mysore</td>
<td><a href="http://www.vidyanidhi.org.in">http://www.vidyanidhi.org.in</a></td>
<td>Theses and dissertations</td>
<td></td>
<td>DSpace</td>
<td></td>
</tr>
<tr>
<td>Archives of Indian Labour</td>
<td>VV Giri National Labour Institute and Association of Indian Labour Historians</td>
<td><a href="http://www.indialabourarchives.org">http://www.indialabourarchives.org</a></td>
<td>Historical documents</td>
<td></td>
<td>sets of collection Greenstone</td>
<td></td>
</tr>
</tbody>
</table>

**FINDINGS**

It has been already pointed out that the institutional repositories in India are using open source software like Greenstone Digital Library Software, DSpace, GNU EPrints and so on. Using Greenstone, chapter-wise or section-wise representation of document is possible. Using DSpace or GNU Eprints software,
self-archiving of digital documents is possible after usual peer review process. This saves the time of information professionals for metadata creation. These software efficiently organize metadata of documents, maintaining international metadata standards. Metadata helps in representing a document and later on helps in retrieving a document from the database. Institutional repositories in India mostly provide an interface of browsing the collection subject-wise, title-wise, document type-wise and so on. These also provide simple and sometimes advanced search options with single or multiply search criteria for precise retrieval. Some repositories permit users to view and download full-text documents (e.g. EPrints@IISc), whereas some others permit users to view metadata and abstracts only (e.g. ETD@IITK).

In the case of some repositories document collection is growing rapidly (e.g. EPrints@IITD). On the other hand in certain cases document collection is either slowing down or remaining static (e.g. Archives of Indian Labours).

A few subject-specific institutional repositories are there where scholars of a particular subject can contribute their documents on the subject. Librarian’s Digital Library (LDL) of DRTC, Bangalore is a subject-specific repository that provides access to articles, conference papers, dissertations, tutorials and other learning materials pertaining to library and information science. The submission of documents by the professionals from other institutions is permissible in LDL.

A few document type specific institutional repositories also exist where scholars of different subjects can contribute their documents according to the type. Vidyanidhi of Mysore University, Mysore is a document type repository that provides access to theses and dissertations of PhD, M.Phil, M.A., M.Sc, M.Tech, M.E., M.D., M.Com., M.B.A., M.C.A., M.LIS, M.Ed., M.SW, etc. The submission of theses and dissertations by the scholars from other institutions is permissible in Vidyanidhi.

Document types in institutional repositories vary significantly and comprise among others textual, graphical, photographic, audio-visual and multimedia documents. The documents commonly available in Indian institutional repositories are preprints, unpublished as well as peer-reviewed articles, seminar presentations, theses, dissertations, research reports, patents, learning objects, computer files, and so on.

From the past few years, Council for Scientific and Industrial Research (CSIR) is encouraging CSIR research fellows to submit theses both in print format and electronic format (CD-ROMs). Other research scholarship providing agencies are also encouraging. Thus, the academic and R&D institutions are receiving electronic theses and dissertations in various disciplines which can be part of institutional repositories in India. National Chemical Laboratory, Indian Institutes of Technology, Indian Institute of Science, and a few universities have initiated institutional repositories to make these worthwhile research documents accessible throughout the world.

Documents available in institutional repositories are many a time distributed across different divisions or departments within the institution. Figures 1, 2 and 3 depict distribution of documents across various divisions or departments. Figure 1 shows that in IIT Delhi repository Centre for Energy Studies account for the highest number of electronic documents (22.5%), followed by Department of Physics (13.5%). Figure 2 shows that in IISc’s electronic theses and dissertations repository, Division of Electrical Sciences accounted for the highest number of electronic theses (50%) followed by Division of Biological Sciences (17.6%). On the other hand, in IISc’s e-Prints repository, the largest number of documents are from Division of Physical and Mathematical Sciences (26.2%) followed by Division of Biological Sciences (25.4%) [Fig. 3].

The institutional repositories, if available in Internet, can be accessible throughout the world. If an institutional repository offers quality research publications, well-defined metadata elements and user-friendly search interface, it would be popular amongst researchers the world over. Also the role of metadata harvester is crucial here that accelerates access to documents of an institutional repository through various web search engines and search services. Interoperability is to be achieved for platform independent information exchange and harvesting. Figure 4 shows that ePrints of IISc is mostly accessed
by the US-based scholars (39.4%) followed by India-based scholars (7.7%). The data of country-wise access was generated by IISc for top 100 publications accessed during 21 May 2004 to 15 Mar 2005. [2]. During this period total accesses was 122,330 (IISc: 363, Others: 121967), excluding access by web crawlers and metadata harvesters. This figure would encourage other institutional repositories in India and they will have to adopt rational approaches and policies in collection development.

![Figure 1: Distribution of ePrints across Departments in IIT Delhi](image1)

![Figure 2: Distribution of ETD across Divisions in IISc](image2)

![Figure 3: Distribution of ePrints across Divisions in IISc](image3)

![Figure 4: Use of ePrints@IISc across the World](image4)

**Problems of Institutional Repositories in India**

When an institutional repository initiative is undertaken, it is first planned keeping in view the objectives of the respective institution as well as of the initiative. Then a prototype is developed, tested and implemented. After achieving satisfactory results, the information repository is developed and implemented at a higher scale and again tested. Some institutional repositories that have been visited to gather information are not presently accessible through Internet due to various reasons. Some of them seem to be are:
The uniform resource locator (URL) has been changed in due course of time, but other web pages link to the earlier URL (i.e., dead link).

The web server is not operational on 24/7 timeframe that makes it inaccessible to the scholars across the world. To be a globally acclaimed information repository, this should be accessible and operational on 24/7 timeframe.

The service has been discontinued but the web pages that are linked to this IR are not updated.

The web server is down for the time being.

The information professionals responsible for institutional repositories should regularly review their accessibility in order to minimize these problems. They can invite feedbacks from users to provide more user-friendly access to collection.

CONCLUSION

The institutional repositories provide access to research publications and other digital documents of respective institutions. The popularity of this concept is growing rapidly in the higher educational and research institutions to disseminate newly emerged knowledge and expertise. When an institution shares its own knowledge resources that not only accelerates knowledge generation and scholarly communication process, but also increases its visibility across the national and international domains. The collection development of an institutional repository is a crucial point for its success and its long-term sustainability. The collection development policy of institutional repositories needs to be reviewed regularly and users' feedback obtained from the user communities. Some institutional repositories in India have just started its operation. Their popularity can be increased through postings in list servers, web search engines, metadata harvesting services and publicity campaigns. The repositories should be linked from the institutions' respective websites. They may include data on ongoing research projects as well as completed research projects. That will reduce duplication of research efforts by other organizations and also save the precious resources of funding agencies. If the works of an individual researcher get an international attention, his or her publications would be accessed more through institutional repository and would be of great value to its collection. This factor needs to be assessed on regular basis.

REFERENCES


