Institutional Repositories in India: A Comparative Study of Dspace and Eprints

Gouri Sankar Karmakar **Rumi Das** **Juli Thakuria**

This paper identifies and outlines the role & importance of various Institutional Repositories (IR) in India. The rapid advent use of IT has completely revolutionized the library systems and changed library services. In India open access movement is getting momentum in libraries to overcome the access barriers of scholarly publication. The institutional repositories were identified through a study of the literature, as well as internet searching and browsing. Each of the Institutional Repository software has a lot of features, unique facilities and excellent capabilities, which the users could explore and experiment. But in this paper we have tried to compare only, two Institutional Repository software i.e. Dspace & Eprints & its role in building Institutional Repositories along with the comparison the various aspects of D-space & E-print software, the various institutions using it in India for their Institutional Repository. Most of the Institutional Repositories (IR) in India use D-space and E-prints software.

Keywords: Institutional Repository, Open Source Software, DSpace, EPrints, ICT.

1. Introduction

In the age of information, the traditional library concept has been changed. Now, digital library, electronic library concepts are being popularized. Now a day, libraries are not only the store house of printed documents but also the centre of digital information. With the development & application of ICT, the services & the collection of the libraries are globalized. The prime objective of libraries is to keep their clientele up-to-date in their areas of interest. It can be in any format, e.g. preparing bibliographies of the latest literature published in their area of research in a broad way, or simply providing them with the information published in the newspapers, websites, personal blogs etc. on the subject of their interest. Due to information explosion, dwindling budget of the library, space problem, high information demand, increasing fees of journals have forced the libraries to seek other ways by which we can collect store & disseminate information among the users. To solve these types of problems, the concept of IR's are started among the academic institution in India. Now the academic institutions have started to build their own repositories.

2. Institutional Repositories

Institutional repositories are "digital collections that capture and preserve the intellectual output of a single or multi-university community" (Crow, 2002). Institutional repository stores and makes accessible

the educational, research and associated assets of an institution. According to wikipedia, "an Institutional repository is an online locus for collecting, preserving, & disseminating in digital form, the intellectual out put of an institution, particularly a research institution". Some Institutional repositories focus on particular subject domains or documents like a research journal articles, before (preprints) & after (post prints) undergoing peer review, and digital versions of thesis, dissertations etc, but it may also include other digital assets generated by normal academic life, such as administrative documents, course notes, or learning objects. Institutional repositories are partly linked to the notion of a digital library i.e collecting, housing, classifying, cataloguing, preserving & providing access to digital contents analogous with the libraries conventional function of collecting, housing, classifying, preserving & providing access to analogue content. An IR is a digital repository of the research out put of an institution. Institutional repositories are one of the most promising developments that utilize new web technologies to offer a viable & sustainable alternative to the current model of scholarly publishing. Institutional repository is the exhibition of an institution to the world where institution displays it's valuable research program, projects and initiatives to the world. In an academic institution, an IR is created from the research output of the faculty, staff and students & made available to the user both within & outside the institution.

3. Institutional Repository Services

There are a number of reasons for building Institutional repository; Institutional repositories are a new but important area within the educational Landscape. Institutional repositories are a practical, cost effective & statistic means for academic institutions for built partnership with their faculty to advance scholarly communications. Institutional repository provides tools that help faculty, students and researchers disseminate their work to audiences outside the institution. Institutional repository may serve as a complement to traditional forms of publication or as an alternative. Institutional repository enable information seekers to find faculty and student work more easily by organizing and indexing it, making it more visible to colleagues, fund providers and fund providers and employers. Facilitate long-term preservation through centralized planning, support and funding. Demonstrate the significance and relevance of the institutes total research activities. An Institutional repository concentrates the intellectual product created by their students, researchers, faculty members to demonstrate its scientific, social and financial value.

Most prominent Institutional repository establishments reveals that following core services should form an integral part of a typical Institutional repository - Deposit and withdrawal services, access control and rights management, administrative services metadata services, user support and feed back mechanisms, storage space, file naming or name resolution service, search engine, preservation and migration.

4. Open Source Software for Institutional Repository Creation

Open Source Software is defined as computer software for which the source code and certain other rights normally reserved for the copyright holders are provided under a software license that meets the open source definition, that is in the domain . According to www.opensource.org as "open source promotes software reliability and quality by supporting independent peer review and rapid evolutions of source code. To be certified as open source, the license of a programmer must guarantee the right to read, redistribute, modify, & use it freely". The term "Open Source Software" originated as a part of a marketing campaign for free software. A report by Standish Group states that adoption of open source software models has resulted in savings of about $ 60 billion per year to consumers.

The creation of institutional repositories involves the use of suitable software & hardware. For building institutional repositories so many open source software is available today. Different software has different special features. Every details of the each software's are available on the own software web pages and which is searchable & freely usable. To build institutional repository in any institutions, the institution has to select the comfortable one according to their aim and need. Here, we have mentioned some popular institutional repositories software's in details which are being used by different institutions.

Name of Software	Developer	Web address	License	System Requirement	Operating System
ARNO	Academic Research in the Netherlands Online	http://cf.uba.uva.nl/en/projects/arno	IWI	No Specific	
CDSWare	CERN document server	http://cdsware.cern.ch	GNU General Public License	MySQLdatsbase server,&Apache/(PHP, Python)	UNIX
Dspace	MIT Libraries &Hewlett Packard Labs	http://dspace.org	BSD License	Apache web server,Tomcat servlet engine,&postgre SQL relational database system	Unix or Linux
Eprints	University of Southampton	http://software.eprints.org	GNU General Public Licence	Apache,MySQL database,perl language	UNIX
Fedora	Virginia & Cornell Universities	http://www.fedora.info	Mozilla Public License	Sun Java, MySQL,Oracle9	Windows Unix
Greenstone	University of Waikato, UNESCO & the human info(NGO)	www.greenstone.org	GNU General Public License	Apache,MySQL database,perl language	Windows Unix or Linux

iVia	Infomine,Look,MEL& virtual reference library	http://infomine.ucr.edu/ivia/ivia.php	AGPL (13) Free Software License	MySQL,Berkeley DB Management Packages,C++	Linux
Phornesis	CONACYT, ITESM	http://copernico.mty.itesm.mx/phornesis/project	GNU General Public License	No Specific	Unix or Linux
i-Tor	NIWI-KNAW	www.i-Tor.org/en/system_info/about	GNU General Public issue	Java script,MySQL,Jetty Web server	Unix or Linux

5. Institutional Repository in India

There are 490 Institutional repositories using DSpace & 352 Institutional repositories using EPrints in the world. On the other hand out of 52 Institutional repositories in India 23 Institutional repositories using DSpace & 19 Institutional repositories using EPrints (Source: http://roar.eprints.org). Institutional repositories in India are developing rapidly. Open source software, especially DSpace, is increasingly being used for the creation of digital repositories. Now a day, it is a trend of Institutional repository all around. Many of the institution have their repositories which they have built using various open source software. But most of the institutional repositories for testing or trial purpose only and are not maintaining properly. Many of them are closed and are not being updated regularly.

5.1 List of Institutional Repositories using DSpace

Sl No	Name of IR, Institution	URL	Software
1	DSpace@NCRA, IIT, Bombay	http://ncralib.ncra.tifr.res.in:8080/dspace	DSpace
2	National Chemical Laboratory, Pune	http://dspace.ncl.res.in/dspace/index.jsp	DSpace
3	National Institute of Oceanography (CSIR), India	http://drs.nio.org/drs/index.jsp	DSpace
4	Delhi College of Engineering	http://202.141.12.109:8080/dspace/	DSpace
5	Guru Gobind Singh Indraprastha University	http://dspace.ipu.ernet.in:8080/dspace/	DSpace
6	Indian Institute of Technology Delhi research and Electronic Submission of Theses and Dissertations	http://eprint.iitd.ac.in/dspace	DSpace
7	DKR@CDRI, Digital Knowledge Repository of Central Drug Research Institute	http://dkr.cdri.res.in:8080/dspace	DSpace
8	DSpace @ INFLIBNET	http://dspace.inflibnet.ac.in/	DSpace
9	DSpace at Bangalore Management Academy	http://bma.ac.in:8080/dspace	DSpace

10	DSpace at ICFAI Business School (IBS), Ahmedabad	http://202.131.96.59:8080/dspace	DSpace
11	DSpace@University of Hyderabad, IGM Library	http://digilib.uohyd.ernet.in/dspace	DSpace
12	Dspace@IIA, Indian Institute of Astrophysics	http://prints.iiap.res.in/	DSpace
13	Dyuthi at CUSAT , Cochin University of Science & Technology	http://dyuthi.cusat.ac.in/dspace/	DSpace
14	ETD@IISc, Electronic Theses and Dissertations at Indian Institute of Science	http://etd.ncsi.iisc.ernet.in/	DSpace
15	Institutional repository at MDI, (Management Development Institute)	http://dspace.mdi.ac.in/dspace	DSpace
16	ISI Library, Bangalore	http://library.isibang.ac.in:8080/dspace/	DSpace
17	Indira Gandhi Institute of Development Research.	http://oii.igidr.ac.in:8888/dspace	DSpace
18	KR@CIMAP, Central Institute of Medicinal and Aromatic Plants	www.kr.cimap.res.in	DSpace
19	Librarians' Digital Library	https://drtc.isibang.ac.in/	DSpace
20	National Center for Antarctic Research , Goa	http://dspace.ncaor.org:8080/dspace/	DSpace
21	Dspace@nitr, National Institute of Technology.	http://dspace.nitrkl.ac.in/dspace/	DSpace
22	Niscair Online Periodicals Repository (NOPR)	url: http://nopr.niscair.res.in/	DSpace
23	Raman Research Institute Digital Repository	http://dspace.rri.res.in/	DSpace
24	Institute of petroleum management	http://203.77.192.116:8080/dspace/index.jsp	DSpace
25	Sarai Multimedia Digital Archive, Delhi	http://archive.sarai.net/dspace/	DSpace
26	Thapar University, Patiala	http://dspace.tiet.ac.in/dspace/	DSpace
27	Bangalore Management Academy,Bangalore	http://59.92.116.53:8080/dspace	DSpace
28	M. N. Dastur & Company (P) Ltd,Kolkata	http://e-lib.dasturco.in:8080/	DSpace
29	Vidyanidhi Digital Library and E-Scholarship Portal	http://dspace.vidyanidhi.org.	DSpace
30	IIT, Kanpur	http://172.28.64.70:8080/dspace	DSpace
31	eGyankosh, IGNOU, New Delhi	http://www.egyankosh.ac.in/	DSpace

5.2 List of Institutional repositories using EPrints

Sl No	Name of IR, Institution	URL	Software
1	National Centre for Catalysis Research (IIT): Catalysis Database	http://eprints.iisc.ernet.in/	EPrints
2	ePrints@IISc, National Centre for Science Information (NCSI),Indian Institute of Science, Bengaluru	http://eprints.iisc.ernet.in/	EPrints
3	Eprint@DU, University of Delhi	http://eprints.du.ac.in/	EPrints
4	ePrints@IIMK, Indian Institute of Management Kozhikode Scholarship Repository	http://eprints.iimk.ac.in/	EPrints
5	Indian Institute of Information Technology	http://eprints.iiita.ac.in/	EPrints
6	OneWorld South Asia Open Archive Initiative	http://open.ekduniya.net/	EPrints
7	OpenMED@NIC	http://openmed.nic.in/	EPrints
8	ePrints@Catalysis National Centre for Catalysis Research (NCCR)	http://203.199.213.48/	EPrints
9	National Aerospace Laboratories(NAL) Institutional Repository	http://nal-ir.nal.res.in/	EPrints
10	Institute of Minerals and Materials Technology	http://eprints.immt.res.in/	EPrints
11	Eprints@MDRF, Dr. Mohan's Diabetes Specialities Centre,Diabetes	http://mdrf-eprints.in/	EPrints
12	Eprints@mkuMadurai Kamaraj University	http://eprints.mkuniversity.in/	EPrints
13	National Metallurgical Laboratory	http://eprints.nmlindia.org/	EPrints
14	Eprints@SBT MKUSchool of Biotechnology, Madurai Kamaraj University	http://eprints.bicmku.in/	EPrints
15	Eprints@IARI, Indian Agricultural Research Institute	http://eprints.iari.res.in/	EPrints
16	Madurai Kamaraj University Repository	http://eprints.mkuoa.in/	EPrints
17	MedknowEprints	http://eprints.medknow.com/	EPrints
18	CSIR Unit for Research and Development of Information Products, Pune	http://eprints.csirexplorations.com/	EPrints
19	Indian academy of science	http://smart.ncsi.iisc.ernet.in	EPrints
20	Bangalore university	http://202.141.128.119/	EPrints
21	dspace@IIMK, Indian Institute of Management Kozhikode Scholarship Repository	http://dspace.iimk.ac.in/	EPrints

6. Dspace Software

D-space is a digital repository system that captures, stores, indexes, distributes & preserves an organization's research data. D-space is the software of choice for academics, non-profit, and commercial organization building open digital repositories. It is free and easy to install and completely customizable to fit the needs of any organizations. D-space is jointly developed by MIT libraries & Hewlett-Packard labs released in April, 2004.

D-space integrates a user community orientation into a system's structure. This design supports the participation of the schools, departments, research centers and other units typical of a large research institution. D-space was developed in responds to expressed faculty need for an easy to use, dependable service that could manage, host, preserve & distribute materials in any type of digital medium formats i.e. Journal papers, Data sets, Electronic theses, Reports, Conference posters, Videos, Images, jpeg, mpeg, tiff files etc.

Some most important features of DSpace are as follows.

a) DSpace indexes digital content, so users can search and retrieve results quickly.
b) DSpace distributes digital content over the World Wide Web and also searchable through search engines.
c) Easy to upgrade.
d) DSpace preserves your digital materials over the long term. Storing examples of students' projects
e) Storing reusable teaching materials that can use with course management systems.
f) Showcasing students' theses or research works.
g) Keeping track of a publications/bibliography
h) Having a persistent network identifier for work, that never changes or breaks
i) No more page charges for images. You can point to your images' persistent identifiers in your published articles.

DSpace's most important functions are to facilitate preservation and access to digital objects. DSpace has a number of preservation features including the ability for libraries to set preservation support by file type, checksums to ensure file authenticity, and persistent identifiers. DSpace uses open standards to facilitate interoperability and hence makes it easy to re-use metadata and for search services, such as Google, to crawl content.

7. E-prints

E-prints is a free open source software and developed at the University of Southampton and released under GPS license in 2000. E-prints is the most flexible platform for building high quality, high value

repositories, recognized as the easiest & fastest ray to set up repositories of research literature, scientific data, students thesis, project reports multimedia artifacts, teaching materials, scholarly collections, digitized records exhibitions & performances.

E-print is open source software for building open access repositories that are complaint with the open archive initiatives protocol for metadata harvesting. It shares many of the features commonly seen in document management systems but is primarily used for Institutional Repository & scientific journals.

Some of the very useful features of GNU EPrints.org are

a) browsing of records by the "keywords" metadata field & alphabetized listing of the author names;
b) Customization of repository homepage, metadata fields, and document types.
c) It is Unicode-compliant, so the repositories' content and user interface can be in any language that has a script.
d) Depending on the OS on which GNU EPrints is being installed, most of the dependent software can be installed along with the operating system
e) The document types and their corresponding metadata elements can be easily configured, if required.
f) Browse views can be generated for any of the metadata fields
g) The look and feel of the repository site is easily customizable as required.
h) If required, access to full texts can be restricted. At times, to comply with the publishers' policy the repository administrator has to enforce this restriction. The restriction can be revoked anytime. In the latest version of the software, it is possible to send an e-mail request to the authors to get a copy of such restricted articles.
i) Local customizations can be retained when the repository is upgraded to newer versions of the software.
j) Porting the contents of a repository to a physically different server can be done easily.
k) Batch inputting of records can be easily done using the "import_eprints" script.
l) Rendering of mathematical and chemical equations and symbols in their actual form can be achieved by encoding the equations and symbols in Latex.

Apart from the above-mentioned features, E-prints offer different types of services i.e. Hosting & maintaining the repositories, Customizing the repository to user's specification, Importing your legacy data, Providing customized training for repository managers & IT personnel, Assisting with advocacy & promotion, Providing expert project management from project proposed to product launch etc.

The latest version EPrints has the some major advange of less typing, more quality Inaccurate or misspelled data and missing details is a major challenge for repositories, helps users enter quality metadata

with fewer keystrokes by using a name authority. For starters, the repository itself can be used as an authority; metadata already entered in other records - author names, journals, conferences, funding bodies, institutions - is used to create a shortlist. Alternatively, a repository administrator can create a name authority file, or connect EPrints directly to an external authority service.

EPrints introduces a flexible plugin architecture which makes it easy to create new import and export plugins, opening up EPrints to be used with an enormous range of software and services, and making EPrints the most interoperable platform available for building repositories.

8. Technical Specification DSpace & E-Prints

		Dspace	Eprints
1.0	**Standard Information**		
1.1	AOI-PMH version supported	OAI-PMH v.2.0	OAI-PMH v.2.0
1.2	Open source license	BDS	GNU GPL
1.3	Latest version release date	April- 2009	May-2009
1.4	Latest version number	1.5.2	V 3.1.3
2.0	**Software**		
2.1	Operating system	UNIX\MacOSX\Windows	Linux/Unix/ Windows/ platform
2.2	Programming language	Java	Perl
2.3	Database	PostgreSQL/Oracle	MySQL
2.4	Web server	Apache\Tomcat	Apache 1.3
2.5	Search engine	Lucene	N\A
2.6	Client supported Netscape,Mozilla,IE,Lynx	All web browsers	
3.0	**Install base**		
3.1	Number of installation	490	352
3.2	Geographic coverage	Worldwide	Worldwide
4.0	**User Registration, authentication& Password Administration**		
4.1	Password Administration	Yes	Yes
4.2	User Registration Verification	E-mail/X.509	MySQL
4.3	Multiple Authentication Method	Yes	No

5. Content Import/Export

5.1	Upload Compressed File	Yes	Yes
5.2	Upload from existing URL	No	Yes
5.3	Volume import for metadata	Yes	Yes

6. Metadata

6.1	Metadata schema Supported	Qualified Dublin Core	Dublin Core
6.2	Metadata review support	Yes	Accept,Edit,Delete
6.3	Add/Delete Metadata Fields	Yes	Yes
6.4	Support Unicode character set metadata	Yes	Yes

7. User Interface

7.1	Modify interface	Yes	Yes
7.2	Support multiple language interface	Yes	Yes
7.3	Discussion forum support	No	Yes
7.4	Indexed by Google /other search engine	Yes	Possible

8. Search Capability

8.1	Full Text	Yes	No
8.1.1	Boolean Logic	No	No
8.2	Search all descriptive Metadata	Yes	Yes
8.2.1	Boolean Logic	Yes	No
8.3	Search selected metadata fields	Yes	Yes
8.4	Browse	Yes	Yes
8.4.1	By Author	Yes	Yes
8.4.2	By Title	Yes	Yes
8.4.3	By Issue date	Yes	Yes
8.4.4	By Subject term	No	Yes

9. Comparative study of Dspace & E-prints

- Preservation: D-space provides long-term physical storage & management of digital items in a secure, professionally managed repository including standard operating procedures such as back up, refreshing media & disaster recovery where as E-prints is not for long term preservation but for provide web access to materials.
- Operating system: E-prints has been successfully run under Linux, Solaris & Mac OS/X operating system but D-space is run under Unix & OS/X.

- License: D-space based on BSD license system and E-print is released under a GPL license.
- Metadata standard: The D-space submission process allows the description of each item using a qualified version of the Dublin core metadata schema and any metadata schema can be used with E-prints. It is depend on the administrator to decide what type of material will be stored and for that material what type of metadata is best.
- Use in OAI: E-prints is the first professional software platform for building high quality OAI-component repositories. But D-space hasn't connected with OAI.
- File format: D-space software manage a large number of file format viz. pdf, word documents, JPEG, MPEG, TIFF, files. With the use of E-prints software not only PDF can be created but also conversion is possible from pdf to TIFF, JPG, GIF & DOC etc.
- System Requirement: The required system for D Space software is Apache web server, Tomcat servlet engine & postgre SQL relational database system. But Eprints system requirement is Apache, MySQL database, perl language.

Due to the growing number of DSpace users, Hewlett-Packard and Massachusetts Institute of Technology have formed an independent, non-profit organization. DSpace Foundation. The foundation provides leadership and support to the user community and submission process. But when there is a need for technical support and training in using the software, DSpace is suitable. In the world majority of Institutional Repositories are using DSpace as it has tremendous potential and can support numerous forms and formats and supports community-based content policies.

10. Conclusion

Institutional repository is the most powerful tool to publish & provide the efficient service among the community of institutions. In fact, Institutional repositories are sometimes referred to as open digital libraries, and open models, such as open archives, have emerged at every level of intellectual property sharing. University Grants Commission (UGC) also realizing the importance of hosting research activity of the institute, they have forced the institutions to create and develop their own institutional repository. Similarly, government agencies like ICMR, CSIR, DBT & ICAR etc. must make it compulsory for the institute to create Institutional repository. Already INFLIBNET centre has created Institutional repository by using D-Space. Institutional repositories are one of the most promising developments that utilize new web technologies to offer a viable and sustainable alternative to the current model of scholarly publishing. The repositories also serve as a comprehensive publications database of the parent organization, which in turn facilitate better management of research knowledge, better visibility and wider access, better impact and citations, rapid communication of research, long-term preservation. We are mentioned only the most important aspects of IR as it is not possible to find out all difference between D-space & E prints in such a small paper.

References

1. BHAI, Rakesh Kumar. Creation of Digital library using OSS: A comparative study of GSDL, D-space & E-prints. *In: Modernization of Libraries: A Challenge in Digital Era*. Pp 28-46. edited by Satpathy, S K, Bhubaneswar. 2008.
2. Crow, R. The Case for Institutional Repositories: A SPARC Position Paper. Available at www.arl.org/sparc/bm~doc/ir_final_release_102.pdf *(*Accessed on 07/11/2009)
3. Open Access and Institutional Repositories with EPrints. Available at http://www.eprints.org/ (Accessed on 14/11/2009).
4. Open Source Initiative. Available at http://www.opensource.org (Accessed on 08/11/2009)
5. Publishing and Academic Resources Coalition, Washington, DC, available at: www.arl.org/sparc/IR/IR_Final_Release_102.pdf *(*Accessed on 07/11/2009).
6. Registry of Open Access Repositories (ROAR). Available at http://roar.eprints.org/index.php (Accessed on 07/11/2009)
7. RICHARD K Johnson. Institutional Repositories: Partnering with faculty to enhance scholarly communication. *In: DLib Magazine*. Nov. 2002 Vol. 8 No. 11.pp

About Authors

Mr. Gouri Sankar Karmakar, Librarian, Barbhag College, Nalbari and a research scholar of DLIS, Gauhati University.
E-mail: gourisankarkarmakar@gmail.com

Ms. Rumi Das, Librarian, T H B College Jamugurihat.
E-mail: rumi.d8@gmail.com

Ms. Juli Thakuria, Librarian, Dr. B.K.B. College, Puranigudam, Nagaon, Assam.
E-mail: julithakuria@gmail.com