THE ROLE OF OPEN SOURCE SOFTWARE IN BUILDING INSTITUTIONAL REPOSITORY

Dipen Deka

Abstract

Advances in Information Communication Technology (ICT) has created immense methods for creating, storing, maintaining, accessing and preserving the traditional printed documents in digital form. The different publishers have taken the full advantage of publishing the research outputs of the academicians and deprive the institutions and the community of the institution from the research outputs. This paper explores the importance of Institutional Repository (IR) and the role of the Open Source Software (OSS) in building the Institutional Repository of any institution. To publish and serve the community of an institution building institutional repositories is the most feasible solution. We have to take the help of some special software packages to build up an institutional repository and the role of open source software in this regard is very important. The institutions which are economically not strong enough can take the advantage of using open source software to build up their own institutional repository and can expose their knowledge stock to the world.

Keywords: Institutional Repository; Open Access; Open Source Software.

1. Introduction

The services and the collections of the libraries and information centre are becoming global due to the application of Information Communication Technologies. Now information can be accessed from even the remote places also with the help of internet. Due to shrinking budget and the increasing price rate of journals which are the main primary source of information we have to look forward to a new alternative by which we can collect, store, arrange and disseminate information to the consumers. The concept of open access has evolved to find out the solutions for these types of problems. The concept of institutional repositories started many years ago in USA and in several countries of Europe. Now the institutions especially the academic institutions are starting to build their own repository to serve their community locally and globally. In building the institutional repository the institutions can take the help of the open source software which is not only economically feasible but technically well advanced also.

2. Institutional repository (IR)

The concept of capturing and making available the research outputs of an institution and other relevant documents locally to the users of the particular institution through intranet and globally through internet is institutional repository. An institutional repository might include electronic versions

of documents such as research papers, project reports, patents, theses and dissertations. 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. According to C Lynch (2003) institutional repository is "a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members". Institutional Repositories manage and create supporting services to store, preserve, and disseminate an organization's digital information or knowledge assets created by faculty, research staff, and students.

Institutional repository is the exhibition of an institution to the world, where institution displays its valuable research programmes, projects, and initiatives to 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.

The intellectual output of an institution is very diverse and may include the following:
- pre-prints of articles or research reports submitted for publication
- the text of journal articles accepted for publication
- revised texts of published work with comments from academic readers
- conference papers
- teaching materials
- student projects
- doctoral theses and dissertations
- datasets resulting from research projects
- committee papers
- computer software
- works of art
- photographs and video recordings

3. Advantages of setting up an IR

Academic Institutions across the world will gain from a more efficient and cost-effective system of scholarly communication. Establishing an institutional repository also enables a university to publicize its research and teaching programmes by enabling access to the work of its staff and students. The quality of a university's academic output forms an effective advertisement for the institution. Many universities also have record-keeping offices to maintain control over the vast amount of paper produced by committees and departments. An institutional repository can be an effective way of storing and making these documents accessible to authorized users. Although universities may see their own interests best-served by making as much content as possible available on open access, there may be some material to which a university may wish to restrict access to specified groups of users. The material to which a university might restrict access is likely to be material created and intended for internal use or that is not ready for general release.

Academic work available on the internet is read more widely than work published in paper format. Also academic work which is available at little or no cost is read more widely than work published in expensive conventional publications. Depositing academic work in an institutional repository therefore increases the profile of an author on a world-wide basis, increasing both the dissemination and the impact of the research they undertake. Deposit in an institutional repository can also ease - both for the institution and for the academic author - the administrative burden of reporting publications for research assessment and review exercises. Regular submission of an author's work to a repository provides an author with a central archive of their work and a record of publications to add to a cv. Increased access to an author's work can benefit career prospects, and there should be no doubt about it.

4. IR initiative in India

Though in India the IR initiative has started lately, but now a large number of institutions are coming forward to build up their IR. 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, 94 self-archiving of digital documents is possible after usual peer review process. This saves the time of information professionals for metadata creation. This software efficiently organizes 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, whereas some others permit users to view metadata and abstracts only.

Table 1: Few Selected Institutional Repository Initiatives in India

Name of IR Institution	Website address	Software used
DRTC, ISI Bangalore	http://drtc.isibang.ac.in	DSpace
IISc, Bangalore	http://etd.ncsi.iisc.ernet.in	EPrints
IIT, Bombay	http://www.library.iitb.ac	Greenstone
IIT, Delhi	http://eprint.iitd.ac.in	DSpace
INFLIBNET, Ahmedabad	http://dspace.inflibnet.ac.in	DSpace
National Chemical Laboratory	http://dspace.ncl.res.in	DSpace
National Informatics Centre	http://openmed.nic.in	Eprints
NIT, Rourkela	http://dspace.nitrkl.ac.in/dspace	DSpace
IIT, Kharagpur	https://www.dspace.iitkgp.erent.in	DSpace
IIT, Bombay	http://library.iitb.ac.in	Greenstone

5. Open Access (OA)

Open access (OA) is the free online availability of digital content. Open access has become an increasingly important and potentially divisive issue in recent years as journal inflation rates have increased. For many librarians and scholars, journal price inflation is itself the central problem and open access is the solution. Open Access holds promise to remove both price and permission barriers to the scientific communication by using Internet. In fact, 'Open access' (OA) is a step ahead of "Free Access" which removes just the price barriers by providing free access to end users. Open Access removes the permission barrier as well. In other words, under Open Access, the end-user not only has free access to the content but also have the right to further distribute the content.

The salient features of Open Access are

- Open access literature is digital, free of charge and free of copyright
- OA is compatible with copyright, peer review, revenue, print, preservation, prestige, career advancement, indexing and supportive services associated with conventional Scholarly literature.
- OA campaign focuses on the literature that authors give to the world without expectation of payment
- OA is compatible with peer review and all the major OA initiative for scientific and Scholarly literature insist on its importance

There are two primary methods for delivering open access to research literature viz

- Open Access Journals
- Open Access Archives

OA Self-Archiving model is liberal on peer review. It simply provides persistent digital repository where authors / owners of the content may archive their documents (Pre-referred or post-referred). Self archiving can be achieved in three ways viz.

- Putting articles on author web sites
- Depositing articles in disciplinary archives
- Depositing articles in institutional archives and repositories

Out of the above mentioned three methods the third one i.e. IR is the most economical and mostly used now a days.

6.   Open source software (OSS)

The open source is defined at the website www.opensource.org as "Open source promotes software reliability and quality by supporting independent peer review and rapid evaluation of source code. To be certified as open source, the license of a program must guarantee the right to read, redistribute, modify, and use it freely."

Features of the Open Source:

- Open source software is normally created and maintained by developers crossing institutional and national boundaries, collaborating by using internet-based communications and development tools.
- Products are typically a kind of "free", often through a license that specifies that applications and source code are free to use, modify, and redistribute as long as all uses, modifications, and redistributions are similarly licensed.
- Successful applications are tend to be developed more quickly and with better responsiveness to the needs of users who can readily use and evaluate open source applications because they are free.
- Quality, not profit drives the open source developers who take personal pride in seeing their working solutions adopted.
- Intellectual property rights to open source software belong to everyone who helps build it or simply uses it, not just the vendor or institution that created or the software.

Open-source software is computer software whose source code is available under a copyright license that permits users to study, change, and improve the software, and to redistribute it in modified or unmodified form. A source code consists of a set of instructions which are translated into binary codes so that the computer can understand the instructions. With the help of a source code of software a computer function. In case of the closed source programs such as Windows, Oracle etc. the source code is secret to the users. Whenever any problems appear the users have to depend upon the particular software developer to remove the problems.

7.   Open source software used   for building institutional repository

There are so many open source software with the application of which we can build IR. Every OSS has some special features and system requirements. The institute has to select the best OSS to be applied to build the IR of their institute. For checking which OSS will be more suitable for building the IR, they can go directly to the web address of the respective OSS and make a feasibility study, as they can be searched and used freely. Below a table is given where the developers, web address, system requirements and operating system of the different OSS are given.

Table 2. Some popular OSS

Name of software	Developer	Web Address	License	System Requirements	Operating System
CDSWare	CERN Document Server	http://cdsware.cern.ch	GNU General Public License	MySQL database server and Apache / {PHP, Python}	Unix
Dspace	Massachusetts Institute of Technology (MIT) libraries & Hewlett-Packard labs	http://www.dspace.org	BSD license	Apache web server, Tomcat servlet engine and the postgre SQL relational database system	Unix or Linux
E-prints	University of Southampton	http://software.eprints.org/	GNU General Public License	Apache, MySQL database, Perl Language	Unix
Fedora	Virginia and Cornell Universities	http://www.fedora.info	Mozilla Public License	Sun java Software, MySQL/ Oracle 9	Windows Unix
Ganesha digital library	YLTI & IDRC	http://gdl.itb.ac.id	GNU General Public License	Apache web server, MYSQL, Database Perl Lang.	Windows 98/Unix
Green stone	University of Waikato, UNESCO and thehuman Info (NGO)	www.greenstone.org	GNU General Public License	Apache web Server, MYSQL, Database Perl Language	Windows, Linux/Unix
i-Tor	NIWI-KNAW	www.i-Tor.org/en/ system_info/about	GNU General Public Issue	Java script, MySQL Jetly web server	Unix/Linux
iVia	INFOMINE, LOOK,MEL & Virtual reference lib.	http://infomine.ucr.edu/ivia/ivia.php	AGPL (13) Free Software License	MySQL & Berkeley Berkeley DB Management Packages, C++	Linux
Phronesis	CONACYT, ITESM	http://copernico.mty.itesm.mx/phronesis/project	GNU General Public Lic	————	Unix, Linux
ROADS	Inst. of Learn. & Res. Tech. (ILRT) UKOffice of Lib. & Infor. Networking	http://www.ukoln.ac.uk	Artistic License GNU/GPL	HTTP Apache web server & Perl language	Unix

8. Conclusion

Open access movement is growing throughout the world and in the near future our imagination of availability of all information sources at a single point is definitely growing to be true. The institutional repository is no longer a concept. In Europe and USA the different project on IR had already started. In India also slowly setting up of IR has already started, especially by the leading academic institutions like IITs, IISc etc are coming forward. The ministry of HRD and the department of Science and Technology are encouraging in building up the IR of the institutes by providing the necessary fund. The universities of India which are still the major centre of study and research activities should take necessary steps to build their IR to serve their communities and to preserve their valuable treasure of knowledge. The role of UGC and INFLIBNET in this aspect is most important. UGC has already started an initiative to develop a repository for Electronic Thesis and Dissertation (ETD). The universities should digitize their doctoral theses and dissertations and other important product and submit them to the IR of INFLIBNET. INFLIBNET has started the project of building the IR by using the Dspace OSS.

Traditionally, academicians usually publish the research outputs in their preferred print version of journals and presently some of the publisher is also bringing out simultaneously an electronic version of the same. But the publisher concerned in terms of subscription both again restricts the accessibility to such information for the individual as well as for the institution. Here the IR has opened up the most convenient, simple and democratic way to the students as well as the academicians to share and publish their research outputs.

References

1. http://www.opensource.org/
2. http://www.oss4lib.org/
3. http://infomine.ucr.edu/ivia/ivia.php
4. http://sorceforge.net/projects/greenstone/
5. http://www.fedora.info/documentation/
6. http://copernico.mty.itesm.mx/phronesis/project
7. http://software.eprints.org/
8. http://www.fedora.info/documents/WhitePaper/FedoraWhitePaper.pdf
9. http://www.dlib.org/dlib/april03/staples/04staples.html
10. http://www.infolibrarian.com/dlib.htm
11. www.i-tor.org/en/system_info/
12. http://www.ukoln.ac.uk/metadata/roads/what/
13. Johnson, Richard K(2002). Institutional Repositories: Partnering with Faculty to EnhanceScholarly Communication. D-Lib Magazine, November.
14. Lynch, Clifford. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age", ARL Bimonthly Report, 226. At http://www.arl.org/newsltr/226/ir.html.
15. Johnson, Richard K(2002). Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication. D-Lib Magazine, November.