Institutional E-Print Repositories for Schorarly Communication : Issues and Implications

B Maharana   D K Pradhan
B K Choudhury   S K Pathy

Abstract

Institutional e-print repositories offer a strategic response to systematic problems in the existing scholarly journal system and distribution of research output by making faster communication and transformation of scholarly information over the long run. This paper introduces e-print archives in general and institutional repositories in particular. The article also discusses the purpose, architecture, elements, and issues of institutional repositories. A guideline for the design of institutional archive has also been discussed. A detailed list of major institutional archives has been presented.

Keywords : E-print archives, Repositories, Scholarly communication, Open Archive Initiative.

0.   Introduction

In the current networked information environment, individually driven innovation, institutional progress and the disciplinary scholarly practices are shifting dynamically to digital medium. It is the primary duty of the academic institutions that, they would take interest in capturing and preserving the intellectual output of their faculty, students and staffs. Traditionally, the institutional libraries have been serving for preservation of the institutions' intellectual legacy and facilitating the scholarly communication. But now in this digital age, these institutional repositories have changed their model and such repositories serve for providing scholarly communication by accessing research articles, supporting the institutions and libraries, reduce the monopoly of journals by demonstrating the scientific, societal and economic relevance with research activities.

Technological growth, usefulness and its trends developed new efforts in institutional repositories. Online storage costs have dropped significantly which can be afforded by repositories. Standards like Open Archive Initiative Metadata Harvesting Protocol (OAI-MHP), progressive metadata standard like 'Dublin Core' have been extensively used as underlying infrastructure for repositories. The development of free publicly accessible journal articles and extraordinary digital work has led to digital institutional repositories system such as, DSpace (http://dspace.org), California Digital Library (CDL) & e-Scholarship Repository (http://repositories.cdlib.org/), Academic Research in the Netherlands Online (ARNO), Scholarly Publishing and Academic Resource Coalition (SPARC) (http://www.arl.org/sparc/), Dispute (http://dispute.library.uu.nl/), E-print (http://www.eprint.org), and many others. The content of e-print repositories consist of narrowly peer-reviewed journal articles, conference papers, posters, pre-prints, multimedia, dissertations and even primary data. The open access of these materials encourages e-print archives as institutional digital repository for scholarly communication.

1.   What are 'e-prints'?

'E-prints' are electronic copies of academic research papers. They may take the form of 'pre-prints' (papers before they have been refereed) or 'post-prints' (after they have been refereed). They may be journal articles, conference papers, book chapters or any other form of research output. An 'e-print

archive' is simply an online repository of these materials. Typically, an e-print archive is normally made freely available on the web with the aim of ensuring the widest possible dissemination of their contents.

Leslie Chan and Barbara Kirshop (2003), in their article, have categorically classified the open e-print archives into four broad groups: individual archives or self archives, institutional archives, discipline based archives, and other special archives. There are a number of successful discipline based open access e-print archives already in existence. The best known is arXiv (http://www.arxiv.org), a service for high energy physics, mathematics and computer sciences. Another example is CogPrints (http://www.cogprints.soton.ac.uk), which covers cognitive sciences.

These subject-based centralized archives work; but so far they have only been taken up by a limited number of subject communities (Pinfield; 2002). Because of this an alternative model is being suggested by advocates of e-prints: institutional e-print archives. Institutions, are assumed to have the resources to substantiate e-print archives, they also have the organizational and technical infrastructures to support ongoing archive provision. In addition, they have direct interest in wishing to expose their research output to others as this would promote the institution's standing in the research community.

2.   Institutional E-print Repositories

Institutional archives are developed, maintained, and administrated by an organization or scholarly society, commonly by institutions, such as, Universities, R&D establishments, Libraries, Museums, etc. to offer universal e-print access facilities stored in their servers. Scholarly Publishing and Academic Resource-Coalition (SPARC) has defined 'Institutional Repositories' as 'digital collections that capture and preserve the intellectual output of single and multi university community' (Crow; 2002). Similarly, Lynch (2003) is of the opinion that, "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". Hence, Institutional Repositories are an effective organizational commitment for long-tern preservation as well as access or distribution for digital materials and support by a set of Information Technologies.

Institutional repositories are focused on the collection and preservation of all types of research literature, scientific data, learning objects, administrative records, multimedia and any other type of collection (Harnad; 2003). Thus, Institutional e-prints repositories are the globally searchable system of distributed interoperable repositories which will impact on the scholarly communication by facilitating dissemination of research result and the e-print institutional repositories are worked under OAI-PMH umbrella. A growing number of institutions and consortia are actively engaged in setting up and running institutional repositories. A country wise list of open institutional e-print archives could be found in the Appendix-1.

3.   Purpose of Institutional Repositories

The origin of the e-print archives lies in the increasing interest in alternatives to the scholarly publishing paradigms. According to Crow, Institutional Repositories have two main rationales; such as:

- ✍ Scholarly Publishing Paradigm: Institutional Repositories centralize, preserve, and make accessible by institution's intellectual capital and they will form global system of distributed interoperable repositories that will help facilitate reform of scholarly communication system.

- ✍ Institutional Visibility and Prestige: Institutional Repositories serve as indicators of academic quality by capturing, preserving and disseminating the collective intellectual capital. The intellectual product created by the researchers, faculty, and other knowledge workers of an institution, deposited in the Institutional Repository; demonstrates its scientific, social and financial value. Thus, the Institutional Repositories measure institutional productivity and prestige and increased visibility of high quality of scholarship.

4.   Elements of Institutional Repository

An institutional repository is a digital archive of the intellectual product created by the faculty, research staff and students of an institution and accessible to end users both within and outside the institution. In other words, according to Crow (2002), the content of institutional repository carry the following elements:

- ✍ Institutional Defined: Institutional repositories capture the original research and other intellectual property generated by an institution's activity in many fields. In this way, it represents the historical and tangible intellectual assets and output of and institution.

- ✍ Scholarly Content: Depending on the goals of establishment of institution, an institutional repository could contain any work product generated by the institutional faculty, student, non-faculty, researchers, and staff. This material is such as electronic portfolio, teaching materials, annual report, video recording, computer programmed, datasets, photographs and digital materials etc.

- ✍ Cumulative and Perpetual: The role of Institutional Repository for scholarly communication is that the content collected is both cumulative and maintained in perpetuity; in this regard it has two roles:

  (a) In Institutional Repository what ever is deposited is protected under legal right to avoid plagiarism, copy right infringement, etc. to sustain perpetually. Hence, the cumulative nature of institutional repository is scaleable.

  (b) Institutional repository aims to preserve and make accessible digital content on a long-term basis. Digital preservation and long-term access are inextricably linked.

- ✍ Open and Interoperable: The institutional repository must provide access to broader community, user outside the institution must be able to find and retrieve information from the repository, means institutional repository must be open access. Therefore, the institutional repository system must be able to support interoperability in order to provide access with the help of search engines and other discovery tools.

5.   Architecture of E-prints Repositories

The architecture (Fig.1) of e-print service is based on harvesting metadata from OAI-PMH (Open Archive Initiative – Protocol metadata Harvesting) compliant e-print repositories from different institutions, non-institutions or by persons into a centralized database. Once gathered, both the metadata and full text of e-prints will be available or passed to external web server by web supporting protocols and that will be able to enhance metadata records by (Day;):

- ✍ Adding/validating authoritative forms of *author names;*

- ✍ Automatic assigning *subject classification terms;*

- ✍ Analyze the bibliographic reference into structured forms, using the Open URL standards.

This enhanced metadata formed the basis of e-print service. It will be made available to end users in a number of ways:

- ✍ Through general search interface which is integrated with other information gateway;

- ✍ Through the developed configured discovery service by which the academic institutions and other organization are directly embed e-print with their own service.

5.1   Open Archive Initiative

Open Archive Initiative (OAI) is supported by the Digital Library Federation, Coalition for networked information environment. Its mission is to develop and promote interoperability standards that aim to facilitate the effective dissemination of content (Simpson; 2004). The OAI has given momentum to any type of institutional archives that contain e-prints of published journal papers produced in research and education institutions to enhance scholarly communication.

The OAI archives can be disciplinary or institutional. The facilitating software is the OAI-PMH (Open Archive Initiative Protocol for Metadata Harvesting) which creates the framework for interoperability between distributed e-print archive/ repositories servers by enabling metadata format to harvest and aggregate into one searchable database/ interface. The metadata format is based on the 'Dublin core' metadata standard elements (Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights). OAI compliant archive/repositories server may be full text papers or may not and the interoperability is of open (free) accessible or not (Hichcock; 2003).

Fig-1
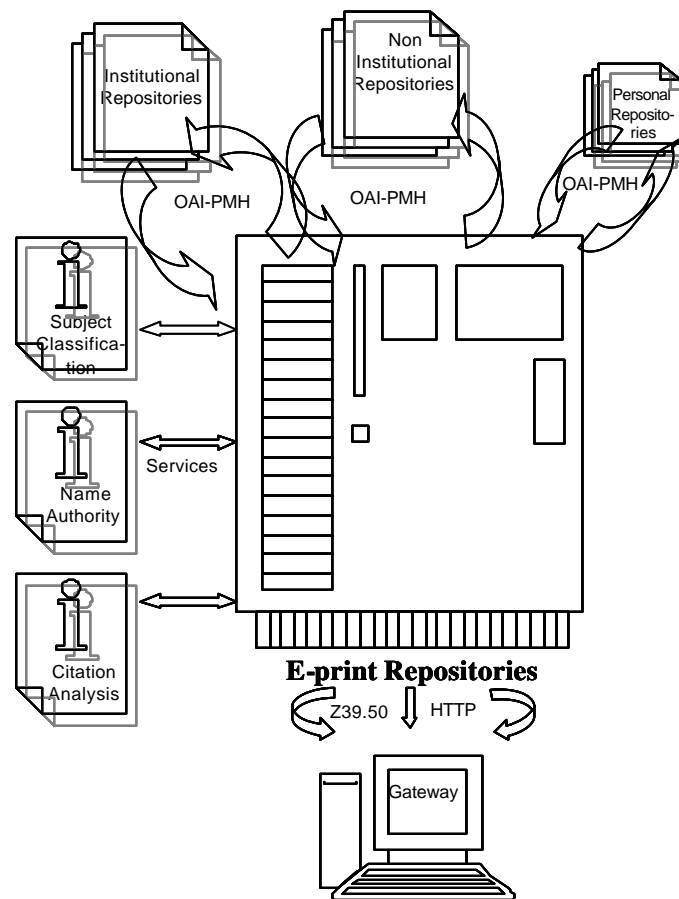
6.   Design of Institutional e-print Repositories

Institutions can provide both the incentives and the means to spread the self- archiving e-print repositories across disciplines by facilitating the following incentives (Harnad; 2003):

- ✍ Installation of 'OAI-complaint e-print archives' using various free software that confirm to OAI – PMH. This will guarantee interoperability of all such e-print archives, as if all the papers deposited there are one seamless global archive, accessible to and navigable by any and every where.

- ✍ Adoption of a policy that all faculties (members) maintain and update standardized online curriculum vitae (CV) for annual review;

- ✍ Mandate that the full digital text of all refereed publications should be deposited in the institutional archive and linked to their entry in the authors' online CV.

- ✍ Train digital librarian/staffs are employed for assisting in self archiving (proxy self-archiving) to the authors who feels that he is personally unable to self-archive for them;

- ✍ Digital librarians, in collaboration with web system staff; provide the proper maintenance, backup, mirroring, upgrading, and preservation of e-print archives (Mirroring and migration are handled in collaboration with that counterparts/institution which supports OAI complaints e-print archives).

7.   Issues and Challenges of e-print Repositories

There are number of issues related to self-archiving institutional e-print repositories, which could be related to a lack of awareness or opportunity, or a less number of records in those repositories, etc , but there are a number of practical or cultural issues which are stated as below (Pinfield:2003):

7.1   Copyright

The main obstacle in the success of institutional e-print repository is the traditional assignment of copyright to the publishers. In general cases when the paper has been accepted for publication then the author has to assign the copy right to the publisher but in case of e-print repository the exclusive license is excluded in the publication of paper.

7.2   Peer-review and quality control

Review is an essential part of the existing scientific and scholarly publishing process. But in this repository the peer-review is outside the scope which can be needed for quality control.

7.3   Long term Preservation

Long term preservation is a potential problem for e-print repositories as digital preservation has always been a challenge for all.

7.4   The popularity of traditional journal

However e-print solves the problem of serial pricing crisis and permission crisis,   but the traditional published journal is more popular in scholarly communication medium because it solves the problem of copy right, quality control, long-term preservation, etc. So the popularity of traditional journal for scholarly communication medium is a partition in the way of e-print repositories.

8.   Conclusion

Institutional repositories are now being recognized as a significant way of valuing and show casing as institution's intellectual assets. It is a major tool in opening access to research. OAI-compliant e-print archives are a real opportunity to improve the access to the research literature to enhance the scholarly communication process. Library and information professionals should have the vision to be leading on this important innovation taking the lead on these important developments. Although the institutional repositories have a huge potential for information communication model, they still need more testing and implementation issues, need to be explored.

Appendix-1
SPARC List of Institutional E-print Archives (Source: http://arl.org/sparc)

Sl. No.	Country	Name of the Repositories	Institution/ Organisation	Contents	System Software
1	AUSTRALIA	EPrint Repository http://www.eprints. anu.edu.au/	Australian National University	Preprint, which have been sent for publication, Post-prints, etc.	Eprint.org
2	CANADA	Papyrus http://papyrus.bib.um	Universitie de Montreal	Preprints, articles, and other research papers	Eprint.org
3	DENMARK	Electronic Library http://www.aub.auc. dk/phd/	Aalborg university	Research papers and publications of lectures and researchers (PDF)	In-house web-based
4	FRANCE	Archive Electronique http://jeannicod.ccs d.cnrs.fr/	Institute Jean Nicod	Preprints, published articles (in journals and anthologies), published correspondence	Eprint.org
5	GERMANY	Eldorado http://eldorado. uni-dortmund.de	Universitat Dortmund	Preprints, published articles(in journals and anthologies), published correspondence	Hyperwave
	GERMANY	MILESS:Die Essener Digital bibliothek http://miless.uni-essen.de/	Universitat Essen	Preprints, published articles, teaching materials, theses & dissertations, & multimedia files	MyCoRe
	GERMANY	Online publications http://elib.uni-stuttgart.de/opus/	Universitat Stuttgart (OPUS)	Preprints, journal articles, proceedings, lecture notes, these & dissertations.	OPUS System
	GERMANY	KOPS-Databank http://www.ub.uni-konstanz.de/kops/	Universitat Konstanz	Preprints, published articles, teaching materials, theses & dissertations	OPUS System
6	INDIA	http://eprints.iisc. ernet.in/	Indian Institute of Science	Preprints, post prints & others scholarly publications.	Eprints.org
7	IRELAND	Eprint archive http://eprints.may.ie/	NUI Maynoot	Preprints & post prints, research papers and other materials	Eprint.org

8	ITALY	Archive E-prints http://e-print.unifi.it/	Universita degli studi de Firenze	Didactic materials, technical reports, theses, working papers preprints as well as published articles conference papers and chapters from books.	Eprint.org
9	THE NATHER LANDS	Digital Academic Repository http://dare.uva.nl/en/	Universiteit van Amsterdam	Scientific publications, reports, preprints, articles, books, chapters, book review, inaugural lecture & dissertations.	ARNO and DLXS
	NATHER LANDS	Electronic documents http://137.120.22.236 /www-edocs/	University of Maastricht	Primary research papers	Eprint.org
	NATHER LANDS	Disputehttp://dispute. library.yy.nl/	Utrecht University	University publications (Full Text) online dissertations.	In house
10	SWEDEN	Electronic Research Archive http://www.hk-r.se/fou/	Blekinge Institute of Technology	Research papers(PDF) web based	In house &
	SWEDEN	Publications http://epubl.luth.se/	Lulea University of Technology	Abstracts describe research papers, technical reports, theses & dissertations(PDF)	In house & web based
	SWEDEN	LUFThttp://www.lub. lu.se/luft/	Lunds Universitet	Teaching materials, report series and research papers.	In house & web based
11	SWITZER LAND	CERN Document Server (CDS)		Over 630,000 bibliogra-phic records, 250,000 full text documents, preprints, articles, books, journals, photographs and many more.	In house
12	UNITED KINGDOM	http://eprints.bath. ac.uk (UKOLN)	University of bath	Preprints and post prints research papers and others research materials.	Eprints.org
	UNITED KINGDOM	Glasgow ePrints Service http://eprints. lib.gla.ac.uk/	University of Glasgow	Full text of the research out put of university scholars, scientist and researchers.	Eprints.org
	UNITED KINGDOM	Nottingham ePrints http://eprints. nottingham.ac.uk/	University of Nottingham	Preprints, post prints and offprint of published papers.	Eprints.org
13	UNITED STATE	eSchlorship http://repositories.cd-lib.org/eschlorship/	California Digital library	Any research scholarly output by university research units, centers, or departments.	Berkeley Electronic Press

US	CODA http://coda.caltech.edu/	Caltech	Completely scholarly research or educational materials in the final form submitted or sponsored by Caltech professional.	Eprints.org
US	DSpace https://hpds1.mit.edu/	MIT	Digital research articles, preprints, technical reports, working papers, conference papers, images and more.	DSpace
US	Hof Prints http://hofprints.hofstar.edu	Hofstar University	Papers written by Hofstar faculty and administrators and papers delivered by Hofstar members and conference paper sponsored by them	Eprints.org
US	Digital Library and Archives http://scholar.lib.vt.edu/DLASPS/	Virginia Polytechnic Institute and State University	Preprint, published . articles, image, theses and dissertation	

9. References

1. Crow ,R, The Case for Institutional repositories: A SPARC Position paper.2002. <http://www.arl.org/sparc/lr/ir.html>

2. Day, M, Prospects for institutional e-print repositories in the United Kingdom. Version 1.0, may 2003. <http://www.ukoln.ac.uk/m.day/>

3. Harnad, S, Eprints: electronic Preprint and Postprints. Encyclopedia of Library and Information science, ed. Marcel Dekker.2003. <http://www.cogsci.soton.ac.uk/~harnad/>

4. Hichcock, S, Metalist Open Access E-print Archives: The Genesis of Institutional Archives and Independent Service.2003. <http://www.arl.org/newsltr/227/ir.htm>

5. Lynch, C. A,. Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age 2003. <http://www.arl.org/newsltr/226/ir.htm>

6. Open Society Institute. A guide to Institutional repository. <http://www.soros.org)/openaccess/software/>

6. Pinfield, S, Open Archive and UK Institution: An Overview. 2003. <http://www.dlib.org/dlib/pinfild/o3pinfild.htm>

7. Pinfield, S, Gardner, M and Maccolla, J,. Setting up an institutional e-print archive.

8. Ariadne. Issue 31; 2002. <http://www.ariadne.ac.uk/issue31/eprint-archives/intro.html>

9. Simpson, P and Hey, J M N, Institutional E-Print Repository for Research Visibility. <http://www.soc.soton.ac.uk>

10. Chan, L and Kirshop, B. Open Archives Opportunities for Developing Countries: Towards a suitable distribution of global knowledge. Ariadne. 30; 2001. <http://www.ariadne.ac.uk/isssue30/oai_chan/intro.html>

11. http://dspace.org

12. http://repositories.cdlib.org

13. http://www.arl.org/sparc/

14. htp://dispute.library.uu.nl

15. http://www.eprint.org

16. http://www.arxiv.org

17. http://www.cogprints.soton.ac.uk

About Authors

Bulu Maharana is working as a Lecturer in the Post Graduate Department of Library & Information Science, Sambalpur University, Jyoti Vihar, Orissa since 2001.   He has a professional experience of working in Indian Institute of Management, Indore for more than two years. He has a good number of publications in LIS journals and presented papers in conferences.
Email : bulu_maharana@yahoo.com

Dibya Kishor Pradhan, presently working as Associate Lecturer in the Post Graduate Department of Library & Information Science, Sambalpur University, Jyoti Vihar, Orissa.
Email pradhandibya1@yahoo.co.in
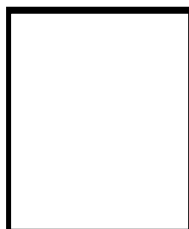
Dr. B. K. Choudhury is currently Head and Coordinator, UGC-DRS-SAP Autonomous Department of Library & Information Science, Sambalpur University, Jyoti Vihar-768019. He is a product of Jadavpur, Karnatak and Utkal University. He has a professional experience of 12 years and teaching for 21 years. He has a good number of publications both in form of journal article and books.
Email : bkc_123@rediffmail.com, bkc_2008@yahoo.co.in
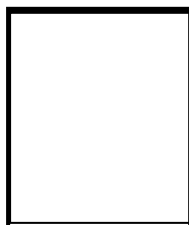
S. K. Pathy is working as Information Scientist in the Prof. B. Behera, Central Library, Sambalpur University, Jyoti Vihar. Earlier to this he was working as Librarian, Reliance School, Jamnagar and also worked with CEE, Ahamadabad. He has more than five years of working in the computerized library environment.
Email : skpathy@rediffmail.com