
INSTITUTIONAL REPOSITORIES TOWARDS KNOWLEDGE MANAGEMENT

Ramesha

S. Gopalakrishnan

Abstract

Institutional Repositories (IRs) are a visible manifestation of the emerging importance of Knowledge Management (KM) in higher education. Paradoxically, "scholarly respect for knowledge and a desire to ensure academic freedom make most institutions reluctant to manage knowledge of any sort". This paper explains the concept of Institutional Repository (IR) and brings out the definitional analysis of IR definitions. Briefly describes the initiatives taken towards building IRs. States the logical Principles of IR and emphasizes its significance in Academic and Research Institutions. Further highlights the benefits in developing IRs. Explains the availability of free and fee software in the design and development of IRs and enumerates features of each software. Finally urges to adopt relevant standards to achieve globalization and access of IRs by universal community.

Keywords: Scholarly Publications, Institutional Repositories, Knowledge Management, Open Source Software, Open Source Standards.

1. Introduction

Today institutions are deploying increasingly complex educational storage and delivery systems ranging from course management systems, personalized portals, student portfolio systems, streaming media services, and Web content management systems. Substantial investments in integrated library systems have already been made, using the underlying IT infrastructure to manage these services. These institutional systems are integrated to various degrees, but do not yet provide the type of seamless integration that can leverage new approaches to teaching, research, and administration. Furthermore, the ability to capture and creatively use rich media objects in the present environment lags behind our ability to capture and use data & still images. Finally, there is a lacking in using a structured approach – both technological and procedural – to capture and preserve the intellectual output of the campus community for future access.

2. What is Institutional Repository?

An institutional repository consists of formally organized and managed collections of digital content generated by faculty, staff, and students at an institution. The content of these repositories can be made available for integration with on-campus library and course management systems. It can also be made available to colleagues and students at other institutions, as well as to the general public. When one uses the term "repositories" today, it is nothing but speaking about one of many different technologies that support the storage and distribution of digital content. They are:

- Collection-based digital repositories managed by library professionals, either stand-alone or aggregated method;
- Course management systems and associated file stores;
- Collections of research data and reports managed by academic departments;
- Student academic portfolio systems;

-
- Institutional file storage systems;
 - Digital asset management workflow systems; or
 - Web content management systems used by institutions or departments to store and stage Web content.

While many of these components can play roles in capturing and managing digital content, an institutional repository is a more specific concept – a centrally managed collection of institutionally-generated digital objects designed to be maintained in perpetuity.

3. Institutional Repositories: Some Definitions

Since institutional repositories are a fairly recent development, it is not surprising that there are different views about what constitutes an institutional repository (IR). While these definitions vary, IRs are fairly easy to recognize. Clifford Lynch (Lynch, 2003) has defined an IR as follows:

In my view, 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. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution.

Mark Ware (Ware, 2004) adds OAI-compliance in his IR definition: An institutional repository (IR) is defined to be a web-based database (repository) of scholarly material which is institutionally defined (as opposed to a subject-based repository); cumulative and perpetual (a collection of record); open and interoperable (e.g. using OAI-compliant software); and thus collects, stores and disseminates (is part of the process of scholarly communication). In addition, most would include long-term preservation of digital materials as a key function of IRs.

4. Institutional Repository Initiatives

There are a number of institutional repository initiatives underway within higher education. Some of the most visible initiatives include:

1. The Open Archives Initiative develops and promotes interoperability standards to facilitate the dissemination of content. Their Metadata Harvesting Protocol provides mechanisms for multiple disciplines to contribute to an institutional repository using common metadata. Participants include the Library of Congress, Harvard, Virginia Tech, Los Alamos, Cornell, CNI, NSF, and the Mellon Foundation.
2. MIT's DSpace is an open source software platform that enables capture and submission of works, distribution of those works, and long-term preservation of assets. DSpace endeavors to create a federated collection of intellectual resources from the world's leading research institutions.
3. Harvard's Digital Repository Service (DRS) provides Harvard owners of digital material with a storage and retrieval system for their collections. Services and facilities include an electronic storage facility, management of administrative and structural metadata, preservation policies and procedures, and delivery of objects to front-end systems such as online catalogs.

-
4. The Flexible Extensible Digital Object and Repository Architecture (Fedora) is a foundation for developing interoperable digital libraries and institutional repositories using XML and Web services. Jointly developed by the University of Virginia and Cornell University, and sponsored by the Mellon Foundation.
 5. The Canada ARL (CARL) Institutional Repositories Pilot Project is implementing institutional repositories at nine Canadian universities.
 6. The University of Southampton's EPrints initiative is designed to manage disciplinary or institutional print collections, rather than digital collections. Eprints software is OAI compliant and freely available under a GNU license, and is in use at California Institute of Technology, the University of Queensland, and other institutions

5. Principles of IR

- Digital content Support and promote creation and preservation of digital research outputs.
- Discovery Support and promote discovery of related digital research outputs across content collections and disciplines.
- Collaboration Support and promote national and international collaboration to enhance management and interoperability of digital content.
- Connection Support and promote connections between digital research, learning and administration services.
- Standards Support and promote use of relevant open standards, such as the OAI (Open Archives Initiative) Protocol for Metadata Harvesting.
- Software Support and promote use of existing systems as proven tools for creating OAI-compliant repositories.
- Diversity Support and promote software diversity within an OAI-compliant framework—one solution does not fit all.
- Evolution Support and promote “learn by doing” and use of demonstrators to build shared experience.
- Authors Support and promote processes that make it easy and desirable for authors to submit their research outputs.
- Cost Support and promote approaches that minimise the costs to institutions of participation and compliance.
- Sustainability Support and promote solutions which can be readily looked after, such as the ability to scale up as usage grows.
- Preservation Support and promote choices for digital object formats that enable long term preservation of and access to content.

6. Importance of IR in Institutions

Students and faculty members increasingly recognize the need to store their intellectual output in the form of personal collections, and to make available the results of their work within and outside the institution. Institutions increasingly recognize the need to develop repositories of intellectual output for long-term archival purposes, and to administer the property rights associated with stored assets.

The rationale for universities and institutions implementing institutional repositories rests on two interrelated propositions one that supports a broad, pan-institutional effort and another that offers direct and immediate benefits to each institution that implements a repository.

6.1 New Scholarly Publishing Paradigm

While institutional repositories centralize, preserve, and make accessible an institution's intellectual capital, at the same time they will form part of a global system of distributed, interoperable repositories that provides the foundation for a new disaggregated model of scholarly publishing. This model unbundles the principal functions of scholarly communication, thus presenting the potential to realize market efficiencies previously hidden by the vertically integrated publishing model that now characterizes academic journal publishing.

Altering the structure of the scholarly publishing model will be neither simple nor immediate. The stakes are high for all the well-entrenched participants in the system—faculty, librarians, and publishers—and the inertia of the traditional publishing paradigm are immense. In the near-term, large journal publishers have both the power and the incentive to maintain the status quo: the prestigious journals they control appear integral to the very structure of academic professional advancement. However, digital publishing and networking technologies harnessed by an increasingly dissatisfied library market—as well as by authors themselves—are now driving fundamental changes to this publishing model at an accelerating pace. And new communications paradigms, especially when constructed by the scholars themselves, can eliminate seemingly insurmountable publisher advantages in relatively short order.

6.2 Institutional Visibility and Prestige

Institutional repositories, by capturing, preserving, and disseminating a university's collective intellectual capital, serves as meaningful indicators of an institution's academic quality. Under the current system of scholarly communication, much of the intellectual output and value of an institution's intellectual property is diffused through thousands of scholarly journals. While faculty publication in these journals reflects positively on the host university, an institutional repository concentrates the intellectual product created by a university's researchers, making it easier to demonstrate its scientific, social and financial value. Thus, institutional repositories complement existing metrics for gauging institutional productivity and prestige. Where this increased visibility reflects a high quality of scholarship, this demonstration of value can translate into tangible benefits, including the funding—from both public and private sources—that derives in part from an institution's status and reputation.

While the main purposes of institutional repositories are to bring together and preserve the intellectual output of a laboratory, department, university, or other entity, the incentives and commitments to change the process of scholarly communication have also begun serving as strong motivators. Computers have been ubiquitous on campuses since the late 1980s. Students and faculty are comfortable with the power of online communication. Faculty teachers and researchers want to archive their own materials and have them available on personal or institutional Web sites, these articles, along with the development of the Internet and more powerful search engines, have enabled people to think in practical terms about the establishment of central facilities for storing, archiving, preserving, and making scholarly and artistic materials available. Repositories may be limited to one field, one department, one institution, or a consortium of several institutions. Collaboration through a consortium reduces costs for each member through resource sharing while expanding access to digital materials.

7. Benefits of Institutional Repositories

- Institutional repositories increase the visibility of Indian researchers and Indian Universities/ Institutions. IRs aggregate the research output of a single universities/institution, raising the profile and status of the research being conducted there.
- Institutional repositories increase the accessibility and impact of Indian research both in India and internationally. IRs can elect to eliminate access restrictions and allow anyone with an Internet connection to access the content within.
- Institutional repositories increase the amount and diversity of scholarly output that is collected and preserved. Traditional collection activities have focused on published material; however, many academic digital objects exist outside of the traditional scholarly publishing system, and can be found on departmental websites and personal computers. IRs represents recognition of the importance of a broader range of scholarly material that is produced by Indian academic researchers.
- Institutional repositories facilitate more timely access to research and scholarship. Traditional scholarly publishing methods often have lengthy lag times from submission to publication. IRs can reduce these lag times by capturing this material in a more timely fashion.

8. Choosing Suitable Software

In choosing the solution best suited to its requirements, an institution may wish to consider and evaluate the following options;

1. A hosted service offers the lowest barrier to entry—a third party service provider hosts the repository on behalf of an institution on a fee for service basis. One option is the ProQuest Digital Commons service.
2. Eprints is a low cost option for a repository primarily aimed at open access to article pre-prints and post-prints, including digital theses. A range of object types can be uploaded, including video, audio, images and zip files. There are 11 Eprints repositories in Australia and the University of Tasmania has offered a consulting service to help New Zealand institutions choosing this option. This option offers easy entry, at the price of limited functionality.
3. DSpace is functionally richer and supports a wide range of object types, including text, sound, images and video. It provides detailed implementation guidelines. The institutional repository at the Australian National University uses DSpace. Institutions for which Eprints is not quite suitable may find DSpace more closely meets their needs, without being unnecessarily complex.
4. Acknowledging the success of the Dutch DAREnet repository system, an institution may wish to consider ARNO as an alternative to DSpace.
5. For an institution seeking a powerful, well-architected repository suitable for very large (1,000,000 plus) collections, and that have the technical expertise to support it, Fedora is a possible solution. Both Monash University and the University of Queensland are building services layers to improve Fedora's usability. Monash has a production repository containing about 70 papers.
6. CDSware may be a suitable alternative for large and complex collections. While it does not appear to be widely-used, it is proven in the rigorous and demanding environment of CERN.
7. For an institution considering multiple discipline-based repositories with a metadata harvesting and resource discovery service, i-Tor may be a suitable option.

9. Institutional Repository Software Providers

The following are some of the more well known software developers/vendors offering Institutional Repository software. This is not an exhaustive list but one might examine these when choosing the system that best suits their needs:

1. Archimede
2. Bepress
3. CDSware
4. CONTENTdm
5. DSpace
6. EPrints
7. Fedora
8. Greenstone
9. Open Repository

The abstracted overviews of the above software are given in Annexure. This list of software platforms is not exhaustive. There is several other software platforms that libraries choose to use. The information in this section is as accurate as possible at the time of writing. Software standards and releases change constantly, so be sure to research current offerings using this information as a starting point.

10. Adopt Relevant Global Standards

Common standards facilitate interoperability and development of services such as resource discovery, authentication and authorisation, or federated searching. Institutions wishing their repositories to form part of the emerging global network of linked repositories need to ensure their repositories can interoperate with others. To achieve interoperability, the following areas (and potentially others) require discussion, elaboration and agreement across participating institutions:

- avoid unnecessary limitations on the types and content or format of digital objects
- adopt a common standard for the unique, persistent identification of digital objects
- digital objects must have accompanying descriptive metadata, conforming at least to some form of Dublin Core
- repositories may expose additional metadata specific to their domain needs the format for metadata must be XML-based
- adopt OAI-PMH as the standard for data exchange between networked repositories
- adopt HTTP as the protocol for linking, so that anyone using any browser can access the repository

The software packages described above support these general requirements, but implementation details may vary.

11. Conclusion

It is evident that the potential of institutional repositories to help foster change within the academy will be significant. The concept of institutional repositories develops in both convergent and divergent ways over the next few years. Not every institution will develop a formally managed institutional repository along the lines of DSpace. But every institution that is utilizing course management systems, library catalog systems, and student portfolio systems will see increased “repository-like” functionality in their products. The open source movement, coupled with greater network collaboration among researchers, should give rise to discipline-specific federated repositories hosted by institutions, research projects, or professional associations.

Institutional repositories are a visible manifestation of the emerging importance of knowledge management within higher education. Paradoxically, “scholarly respect for knowledge and a desire to ensure academic freedom make most institutions reluctant to manage knowledge of any sort.” The long-term impact of institutional repositories is likely to change many of the basic assumptions about how intellectual output is managed by individuals, their colleagues, and the academy, and how research itself is conducted.

12. References

1. Besser, Howars, “The Next Stage: Moving from Isolated Digital Collections to Interoperable Digital Libraries,” *First Monday*, 7 (6), June 2002, URL: http://firstmonday.org/issues/issue7_6/besser/index.html (last access 20.11.2005)
2. Chan, Leslie and Kirsop, Bsrbara. Open Archiving Opportunities for Developing Countries: Towards Equitable Distribution of Global Knowledge. access from <http://www.ariadne.ac.uk/issue30/oai-chan/> (last access 10.11.2005)
3. Gaur, Ramesh C et.al. The Institutional Repositories: An Alternative Model for Scholarly Publishing. In. Road Map to New Generation of Libraries using Emerging Technologies. Ed. By T.A.V. Murthy et.al. INFLIBNET Centre, Ahmedabad, 2004.
4. Flecker, Dale, “Digital Archiving: What Is Involved?” *EDUCAUSE Review*, January/February 2003, URL: <http://www.educause.edu/ir/library/pdf/erm0316.pdf> (last access 10.11.2005)
5. Lynch, Clifford, “Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age,” *portal: Libraries and the Academy*, Baltimore: Johns Hopkins University Press, 3(2), 2003, Pp. 327-336.
6. Lynch, C. “Institutional repositories: essential infrastructure for scholarship in the digital age”, *ARL Bimonthly Report*, 2003, No. 226, available at: www.arl.org/newsltr/226/ir.html. (last access 10.11.2005)
7. Powell, Wayne, and Gill, Chris, “Web Content Management Systems in Higher Education,” *EDUCAUSE Quarterly*, No.2, 2003.
8. Ramesha and Nageswara Rao P. Institutional Repositories: An Overview. In. National Workshop on Recent Trends on Digitization., Ed. by Prof. Ramesh Babu and others, S.V. Central Library and Research Centre and Foundation for Information and Communication(FIC), 2005, Pp.62-66.
9. Ramesha and Kumbar B.D. Institutional Repositories: Enhance the Heritage and Scholarly Publications of Universities and Institutions, *University News*, Vol. 43, No.13, March 21-27, 2005, Pp.9-12

-
10. Townley, Charles, "Will The Academy Learn to Manage Knowledge?" *EDUCAUSE Quarterly*, 2, 2003 .
 11. Shearer, Kathleen. "Institutional Repositories: Towards the Identification of Critical Success Factors". *Canadian Journal of Library and Information Studies*, Vol. 27, No. 3, August 2003.
 12. Shearer, Kathleen. "The CARL Institutional Repositories Pilot Project". *D-Lib Magazine*. 9(3), March 2003 (Online). <http://www.dlib.org/dlib/march03/03inbrief.html#SHEARER> (last access 10.9.2005)
 13. "SPARC Position Paper on Institutional Repositories," August 2002, URL: <http://arl.cni.org/sparc/IR/ir.html>
 - URL: <http://www.creativecommons.org>
 - URL: <http://www.openarchives.org>
 - URL: <http://www.dspace.org>
 - URL: <http://hul.harvard.edu/ois/systems/drs/policyguide.html>
 - URL: <http://www.fedora.info/>
 - URL: <http://www.innovationstrategy.gc.ca/cmb/innovation.nsf/SectorReports/CARL>
 - URL: <http://www.eprints.org>
 - URL: <http://www.sfxit.com>
 - URL: <http://www.theospi.org>
 14. Ware, M. *Institutional Repositories and Scholarly Publishing, Learned Publishing*, 17 (2), 2004, Pp.115-24.

ANNEXURE – I

1. **Archimede** URL: <http://www1.bibl.ulaval.ca/archimede/index.en.html>

Availability : Free, open source software, delivered under the GNU general public licence. Download Archimede software from Source Forge: <http://sourceforge.net/projects/archimede>

Technical support: <http://sourceforge.net/projects/archimede/>

Example site

 - Laval University Library
2. **Bepress** URL: <http://www.bepress.com/repositories.html>

Availability: Commercial software, paid license and subscription fees
Flat rate: between \$8k and ~\$50k per year for unlimited series, unlimited papers, cost relates to size of repository, anticipated usage. *Variable rate:* \$4-5k licence fee and usage fees per content series, per posted paper

Technical support: Available through paid software licence.

Example sites

 - Boston College
 - University of California's eScholarship Repository (<http://repositories.cdlib.org/escholarship/>)
 - Florida State University
 - New England Law Library Consortium

3. CDSware (CERN Document Server Software) URL: <http://cdsware.cern.ch>**Availability**

- Free, open source software distributed under the GNU General Public Licence
- Latest version: CDSware v0.3.3
- Download location: <http://cdsware.cern.ch/download/>

Technical support

- Free email support at cds.support@cern.ch or through mailing list: project-cdsware-users@cern.ch
- Paid technical support is also available.

Example site

- CERN document server: <http://cdsweb.cern.ch/>

At CERN, CDSware manages over 400 collections of data, consisting of over 600,000 bibliographic records, including more than 250,000 full text documents.

4. CONTENTdm™ URL: <http://contentdm.com/>

Availability: Commercial software. Pricing based on collection size.

- Rate between \$7,000 for a maximum of 8,000 stored items and \$40,000 with no limit on stored images.
- Additional cost of \$1,000 - \$6,000 for an annual maintenance agreement fee, which is included in the fee for the first year.
- One day of installation assistance and on-site training is available for \$2,500 plus travel expenses.
- Free 60-day trial available

Technical support: Available through an annual maintenance agreement fee, between \$1,000 - \$6,000. Installation support also available for a fee.

Example sites

Full list of organisations using CONTENTdm at

<http://contentdm.com/customers/customer-list.html> including:

- University of Arizona

5. DSpace URL: <http://www.dspace.org>**Availability**

- Free, open source software jointly developed by MIT and Hewlett Packard Labs.
- Latest version: DSpace 1.2.1
- Distributed through the BSD open source licence
- Download at <http://sourceforge.net/projects/dspace/>

Technical support

- DSpace-Tech mailing list for technical questions, discussions: <http://www.dspace.org/feedback/mailling.html>

Example sites

- Cambridge University
- Cranfield University

6. Eprints URL: <http://software.eprints.org>**Availability**

- Current version: GNU EPrints 2.3.6
- Distributed under the GNU general public licence
- Download software at <http://software.eprints.org/download.php>
- Demo server: <http://software.eprints.org/demo.php>

Technical support

- EPrints-tech mailing list: <http://software.eprints.org/maillist.php>
- General announcements and “underground” discussion list also available at <http://software.eprints.org/maillist.php>.
- EPrints wiki: <http://wiki.eprints.org/w/>

Example sites

Full list of 141 known sites at <http://software.eprints.org/archives.php>, including:

- California Institute of Technology
- CogPrints Cognitive Science Eprint Archive

7. Fedora URL: <http://www.fedora.info/index.shtml>**Availability**

- Free, open source
- Download the current release, Fedora 1.2.1 at <http://www.fedora.info/release/1.2/>

Technical support

- Free online support through mailing list: <https://comm.nsdlib.org/mailman/listinfo/fedora-users>
- Fedora WIKI: <http://www.fedora.info/wiki/bin/view/Fedora/WebHome>

Example sites

- Indiana University
- Kings College, London
- New York University

8. **Greenstone URL:** <http://www.greenstone.org/cgi-bin/library>

Availability

- Free multi-lingual, open source software
- Current version: Greenstone v2.51
- Download location: <http://www.greenstone.org/cgi-bin/library?e=p-enhome-utfZz-8&a=p&p=download>

Technical support

- Online support: <http://www.greenstone.org/cgi-bin/library?e=p-endownload-utfZz-8&a=p&p=support>
- Technical email list: <https://list.scms.waikato.ac.nz/mailman/listinfo/greenstone-devel>

Example sites

Full list at <http://www.greenstone.org/cgi-bin/library>

9. **Open Repository URL:** <http://www.openrepository.com/default.asp>

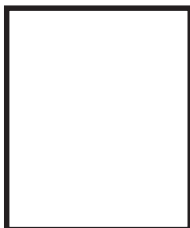
Availability

- Commercial services ranging in price from £5,000 to £10000 for set up and an additional £2,500 to £25,000+ in maintenance fees

Technical support

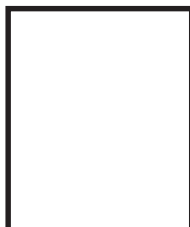
- Full technical support available

About Authors



Dr. Ramesha is working as a Lecturer in Department of Information Science, University of Madras. He has contributed many papers and he is member of several autonomous body.

Email : bramesha@rediffmail.com



Dr. S. Gopalakrishnan is working as Assistant Librarian at Anna University, Chromepet, Chennai

Email :