
INSTITUTIONAL DIGITAL REPOSITORY (IDR) AT CENTRAL LIBRARY, INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR: A CASE STUDY

VISHWANATH PRATAP SINGH

School of Life Science Department of Zoology
Guru Ghasidas Vishwavidyalaya Bilaspur, Chhattisgarh

Abstract

This case study explores the concept of Institutional Digital Repository. It describes various feature and importance of both Institutional Digital Repository and Open Source Digital Library Software named DSpace. It also describes the requirement of hardware and software for setting up the Institutional Digital Repository. It highlights the primary benefits and suggestion for usability of the Institutional Digital Repository.

KEYWORDS :

IDR; Digital Library Software; DSpace; Digital Preservation; Digital-archiving.

1.0 INTRODUCTION:

The intellectual capital of any institutions is represented by the scholarly publications & research material. Institutional Digital Repositories (IDR) are emerging technologies for sharing and management of scholarly publication in academic and research institutions, which can help collect, preserve, store and distribute research material and scholarly publications of faculties, researcher and students in any academic or research institute. Institutional Digital Repositories contain articles, research papers, theses and dissertations, technical report, project report, conference paper and speech/lecture materials, etc., depending upon the purpose for which the repository was created.

There are many Open Source Software's available for creation of the Institutional Digital Repository like DSpace, Fedora, Greenstone, Eprints, Archimede, OPUS and MyCore.

1.1 OBJECTIVES

1. To describe Institutional Digital Repository
2. To Identify the Software in use to create IDR at IIT Kharagpur
3. To describe details of the hardware and software requirement for IDR
4. To describe what can we store in an Institutional Digital Repository
5. To brief the types of materials being included in the IDR of the Central Library IIT-Kharagpur
6. To describe the submitting content in the IDR

1.2 SCOPE

The present study is limited to the Central Library IIT Kharagpur. The study focus on the Institutional Digital Repository by using Open Source Software named DSpace.

1.3 METHODOLOGY

Methodology has tremendous role to execute a study successfully. This study is relating to Institutional Digital Repository at Central Library, IIT Kharagpur. The literature review was limited to published articles covered by Internet resources and online articles. These resources were searched for the entire duration up to January 2013. Multiple search strategies were used on the keywords, 'Digital Libraries', 'Institutional Repository', 'Institutional Repository and 'DSpace', 'Digital Library' and 'DSpace', 'Open Source Software', and 'Digital preservation', . Results of these searches were reviewed to identify relevant articles. This resulted in the 14 relevant articles and various online resources that were used for this study.

1.4 IMPORTANCE

01. A better service to institution's learning community
02. A solution to the problem of preserving Institution's intellectual output
03. A patron friendly Information Retrieval System
04. Contributing to the reform of the entire Institution of scholarly communication and publishing
05. Digital preservation of the documents using IDR
06. Flexible ways to develop existing scholarly communications
07. Multi-patron access facility at the same time
08. Online access for the Institutional Research publications
09. Opportunities for new forms of scholarly communication
10. Reducing patron dependence on Library's print collection

2.0 INSTITUTIONAL DIGITAL REPOSITORY

Institutional digital repositories are emerging technology for collecting, preserving, storing and disseminating of scholarly publication in digital format. This is organized and managed collection of digital content generated by faculty, researcher and students at an institution.

The term “repositories” has one of many different technologies that support the storage and distribution of digital content:

- Collection of scholarly publication managed by academic/research institutions
- Collection-based digital repositories which are managed by information professionals
- Digital asset management workflow systems
- Institutional file storage system
- Web content management systems used by institutions or departments to store, preserve and distributes digital content

While many of these components can play roles in capturing and managing digital content, an institutional digital repository is known as Digital Archives of scholarly publication, which is created by faculty, researcher and students of an institutions or group of institutions accessible to end patrons both within and outside the institutions.

2.1 DEFINITION:

“An Institutional repository is an Organization based set of services which the organization 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” (Clifford Lynch 2003)

2.2 What can we store in a Repository?

The focus of this study is to enable repository services that support storage of,

Academic publications: articles (published articles in journals, magazines and newspapers; peer-reviewed, copyright approved post-print articles; pre-print materials related to published peer-reviewed articles), books, book sections (including conference proceedings and abstracts)

Theses and dissertations: Doctoral theses, Masters theses and dissertations

Grey literature: patents (published only), technical reports, software, project reports, internet publication, documentation and manuals, working and discussion papers, non peer-reviewed conference and workshop materials (posters and speech/lecture materials)

Audio visual items: images, shows/exhibitions, lecture, etc.

3.0 IDENTIFICATION OF IDR AT IIT KHARAGPUR

The Central Library IIT, Kharagpur uses open sources digital library software namely DSpace, developed jointly by MIT Libraries and Hewlett-Packard (HP) in 2002, DSpace is freely available to research institutions worldwide as an open source system that can be customized and extended. DSpace is designed for ease-of-use, with a web-based patron interface that can be customized for institutions and individual department. The IDR of Indian Institute of Technology Kharagpur (IITKGP) collects preserves and makes available in digital format the scholarly output of IIT Kharagpur community. Its interface provides for an easy self-archiving by faculty, and organizes the documents in logical, easily retrieved fashion.

4.1 What is DSpace?

DSpace software is a platform, which serves a variety of digital archiving needs. DSpace is freely available as Open Source Software, which facilities to capture all types of content in digital format. It distributes it over the web. It indexes your work, so patrons can search and retrieve your items. It preserves your digital work over the long time. DSpace provides a way to manage your research materials and publications in a professionally maintained repository to give them greater visibility and accessibility for 24 hour.

4.2 DSpace Submission Workflow

DSpace submission workflow system is a critical part of the DSpace architecture that allows for the submission, processing, and final addition of content to the archives. The system guides the patron through the different steps of the process, which are outlined in Figure 1.

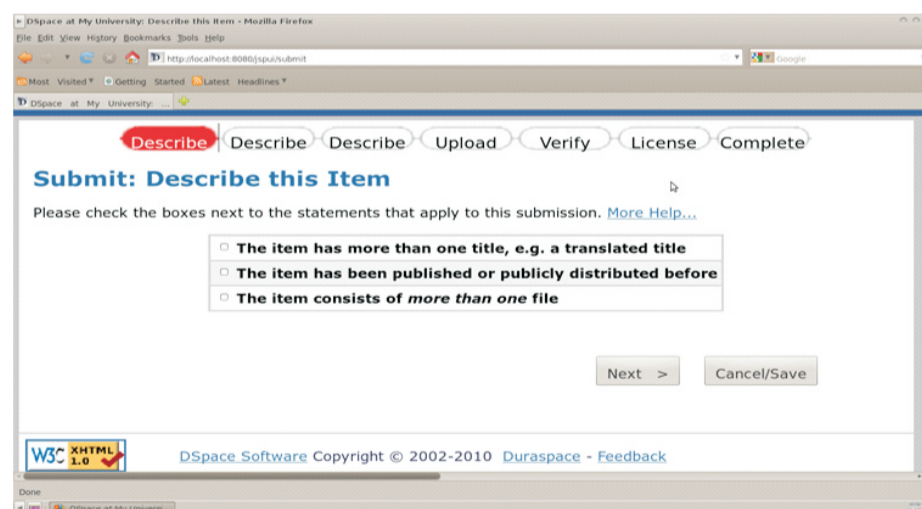


Figure 1- DSpace submission workflow overview

Patron enter metadata about the documents
 The patron select and upload the file on their local system
 Verification of the entered metadata and the file for the submission
 Patron must agree to the license for submission an item in the collection
 The submission process has completed and an item added to achieves

4.3 Features

- 01.DSpace supports OAI – PMH (Open Archives Initiatives Protocol for Metadata Harvesting) Resumption tokens
- 02.DSpace also includes batch tools to import and export items in a simple directory structure
- 03.DSpace exposes the Dublin Core metadata for items that are publicly (anonymously) accessible
- 04.DSpace uses the CNRI Handle System for creating identifiers
- 05.DSpace supports uploading and downloading of bit streams as-is. This is fine for the majority of commonly used file formats such as PDFs, Microsoft Word documents, spreadsheets
- 06.Document discovery and retrieval system
- 07.Digital preservation of all types of documents in digital format

4.4 Digital Preservation Using DSpace

- 01.DSpace provides long-term physical storage and management of digital items in a secure operating procedures such as backup, mirroring, refreshing media, and disaster recovery
- 02.DSpace assigns a persistent identifier to each contributed item to ensure it is retrievable far into the future
- 03.DSpace provides a mechanism for advising content contributors of the preservation support levels they can expect for the files they submit

4.5 Reasons to Use DSpace

Largest community of patrons and developers worldwide
 Free open source software
 Completely customizable to fit our needs
 Used by many educational, government, private and commercial institutions
 Can manage and preserve all types of digital content

4.6 DSpace patrons may browse the contents of IDR in the following ways:-

DSpace administrators organize DSpace items into collections, and include collections in communities. Patrons can browse the structure of communities and collections, in an outline view. Each outline entry includes text describing the contents of the corresponding community or collection. From the DSpace home page, patrons can browse all items in DSpace by title, author, or issue date. From a community or collection home page, patrons can initiate a bounded browse within that community or collection. DSpace supports bounded browse by title, author, or issue date.



Figure 2- Community, collections and Item

5.0 PRESENT STATUS OF IDR AT IIT KHARAGPUR

Central Library of IIT Kharagpur is the first IIT Library to setup the IDR in which the digital documents are organized in to three major categories of collection, such as Administration, Library and Departments / Centers / Schools under the common heading as "Community & collection". Thus collection against each of the above mentioned Community reflects the core area of activity of the community concerned as its main subject of collection. New communities will be created as and when the same is required to give berth to new and emerging subject areas and activities in this Institute. At present 713 PhD theses and almost all the UG Examinations question papers of IIT Kharagpur are available from this IDR.

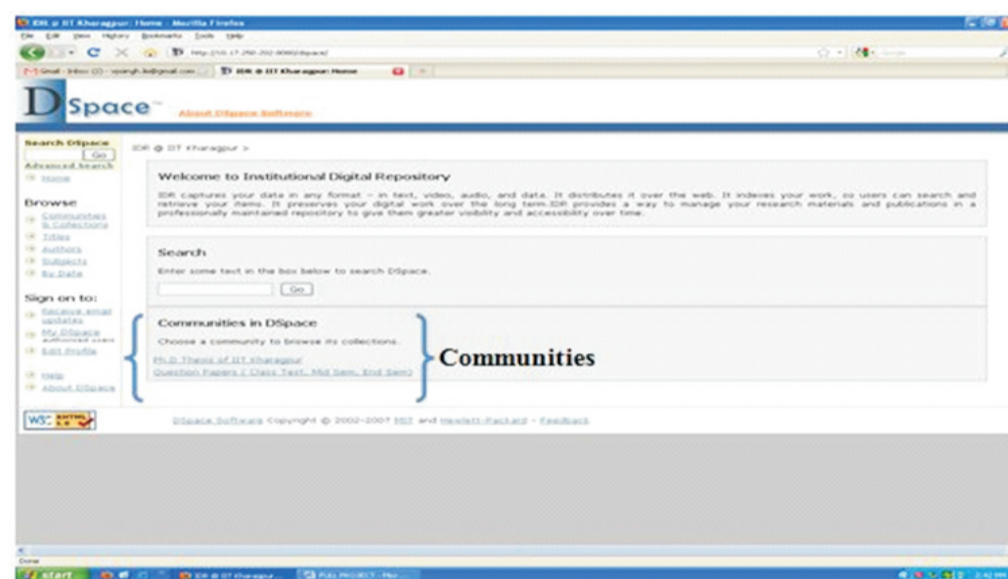


Figure 3- Main page of Institutional Digital Repository at IIT Kharagpur with communities

Central Library IIT, Kharagpur have included all the departments / centers and administration as individual community. Each community has two distinct class of collection. They are:

PhD Theses
Question Papers

6.0 SYSTEM REQUIREMENT

There are two types of requirement for the implementation of Institutional Digital Repository. Those are hardware and software requirements.

6.1 Hardware Requirement:

01. Any computer that can run UNIX type OS (Server Configuration is preferred)
02. RAM memory about 2GB
03. The repository size will decide the storage capacity (hard disk size)
04. High-Speed Face-up Book Scanner with OCR Software (to digitize the print documents)

High-Speed Face-up Book Scanner with OCR Software (to digitize the print documents)

Type	Face-up scanner with Maximum, A2 (17" X 23") size.
Image sensor	7500 pixels reduction type line CCD
Method of scanning	Scanning through Mirror rotation.
Original thickness	Max. Thickness: A2 – 50mm (2") / A3 – 100mm (4") books, files.
Resolution	200 / 300 / 400 / 600 dpi.
Scanning size	Spread: A2, A3, A4, A5, B3, B4, B5, Single: A3, A4, A5, A6, B4, B5, B6
Scanning speed	Less than 4.5 seconds (A4 size lateral document, 400 dpi, and book mode).
Exposure / contrast control	Automatic and Manual (9 adjustment levels each).
Interface	SCSI – 2 / Video Interface.

Table 1: Specification of the scanner**6.2 Software Requirement**

01.DSpace 1.4.2 or latter

Download: <http://dspace.org/>**6.2.1 Supporting Other Software to Install and Run the DSpace Software**

01.UNIX-like OS (Linux, HP/UX etc)

02.Java 1.5.0 (JDK 1.5.0_1

4) Download: <http://java.sun.com/javase/downloads/index.jsp>03.Apache Ant 1.5 or later (Java make-like tool) Download: <http://ant.apache.org/>04.Apache Tomcat Download: <http://tomcat.apache.org/whichversion.html>05.PostgreSQL 8.2 or later, an open source relational database. Download: <http://www.postgresql.org/>**7.0 BENEFITS OF IDR AT IIT KHARAGPUR**

There are many benefits of the IDR at IIT Kharagpur such as

01.A repository enables the institute to publish its own scientific research and to make it available to all of its researchers

02.It can be access 24 hours a day

03.It has a safe, backed-up and secure place to store institution scholarly works

04.Make possible easy access to Electronic Theses and Dissertations (ETD). However only print version of theses, dissertations were available a few years back

05.More than one person can access a particular document at the same time

06.Preserves and provides long-term access to the scholars' research output

07.Provide new opportunities for the archiving and preservation of valuable digital works

08.Provides a global platform for local research and hence improved visibility

09.Reduce duplication of records and inconsistencies in multiple instances of the same works

10.The quality of the institute's intellectual output can be disseminated effectively and efficiently

11.The visibility of the academic output of Indian Institute of Technology will be increased

CONCLUSION

The IIT Kharagpur has successfully implemented an IDR using the Open Source Software (DSpace). Central Library IIT Kharagpur, collects, preserves and disseminates in digital format the research output (PhD theses and Questions paper.) of IIT Kharagpur research community. It enables the Institute community to deposit (self archiving) their preprints, postprints and other scholarly publications using a web interface, and organize these publications for easy retrieval. At present, the access of 'Institutional Digital Repository' is restricted within the IIT Kharagpur campus LAN only and submission of documents to this repository is also limited to the IIT Kharagpur research community

only. It provides a central platform for local research and hence improved visibility. There are some challenges on the road that leads to the kingdom of IDR technology. They are using the Open Source Software named DSpace for the creation of Institutional Digital Repository. Actually the commercial support for the software exists, neither for installation nor for later technical issues.

REFERENCES

1. Bailey Jr, C. W. (2008). Institutional repositories, tout de suite. available on-line at <http://www.digital-scholarship.org/ts/irtoutsuite.pdf>.
2. Barton, M. R., & Waters, M. M. (2004). Creating an institutional repository: LEADIRS workbook.
3. Biswas, G., & Paul, D. (2010). An evaluative study on the open source digital library softwares for institutional repository: Special reference to Dspace and greenstone digital library. *International Journal of Library and Information Science*, 2(1), 1-10.
4. Crow, R. (2002). The case for institutional repositories: a SPARC position paper. *ARL Bimonthly Report* 223.
5. Crow, R. (2004). *A guide to institutional repository software*. New York: Open Society Institute.
6. Das, B. (2008). *Development of a Digital Library using DSpace Open Source Platform* (Doctoral dissertation, Indian Institute of Technology).
7. Hockx-Yu, H. (2006). Digital preservation in the context of institutional repositories. *Program: electronic library and information systems*, 40(3), 232-243.
8. Johnson, Richard K. (2002) *Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication*. *D-Lib Magazine*.
9. Kelly, J. C. (2007). Creating an institutional repository at a challenged institution. *OCLC Systems & Services*, 23(2), 142-147.
10. Lynch, C. A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *portal: Libraries and the Academy*, 3(2), 327-336.
11. Smith, M., Barton, M., Bass, M., Branschofsky, M., McClellan, G., Stuve, D., & Walker, J. H. (2003). *DSpace: An open source dynamic digital repository*.
12. Sutradhar, B. (2006). Design and development of an institutional repository at the Indian Institute of Technology Kharagpur. *Program: electronic library and information systems*, 40(3), 244-255.
13. Thakuria, J. (2008). *Building An Institutional Repository With DSpace*.
14. Yeates, R. (2003). Institutional repositories. *Vine*, 33(2), 96-101.
15. DSpace at IIT Kharagpur. <http://www.library.iitkgp.ernet.in/>
16. DSpace. <http://dspace.org/>
17. Institutional repository. http://en.wikipedia.org/wiki/Institutional_repository
18. DSpace. <http://en.wikipedia.org/wiki/DSPACE>
19. Open Source Software. http://en.wikipedia.org/wiki/Open_source_software