**Design, Development and Validation of Open Source Platform for AIISH Digital Repository and Online Public Access Catalogue**

A non-funded Research Report

**Principal Investigator**

Dr. Shijith Kumar, C

**Co-Investigators**

Mr. Nanjunda Swamy, M.,

Mr. Raghavendra, G.N.

Mr. Lokesh, P

Mr. Nidheesh David Kuruvila

**2021**

Library and Information Centre

**All India Institute of Speech and Hearing**

**Mysuru**

**SUMMARY**

**Project Summary**

The All India Institute of Speech and Hearing is one of the early implementers of the Institutional Repository (IR) and Integrated Library Management System (ILMS) among the higher educational Institutions in the country. However, both the solutions were developed on proprietary software platforms with limited functionalities and are being continued even today without any further modification or addition of features. On the other hand, the IR and ILMS fields at present are dominated by freely available open source applications with advanced features and functionalities. This project designed and developed an open source-based IR and Online Public Access Catalogue, the end-user component of the ILMS. The new platform serves as a single-point access to all the resources and services of the Library and Information Centre, All India Institute of Speech and Hearing, Mysuru.

**Acknowledgements**

We sincerely acknowledge Dr. M. Pushpavathi, Director, All India Institute of Speech and Hearing, Mysore for her encouragement and support towards the completion of the project. We are grateful to Dr. S.R. Savithri, former Director for sanctioning the project and we thank her for her confidence in us. Thanks are also due to Mr. Mahendra, Clerical Assistant, Ms. Bhanumathi, Clerical Assistant and all our colleagues in the Library and Information Centre, AIISH, Mysore.

.

**Investigators**

**Contents**

**Preliminary Section Pages**

1. Project Summary
2. Acknowledgement
3. Contents

**Main Section**

**Chapter 1 Introduction**

1.0 Institutional repositories

1.2 Types of contents in Institutional repositories

1.3 Importance of Institutional Repositories

1.4 Statement of the Problem

1.5 Objectives of Study

1.6 Significance of the Study

1.7 Study Methodology

1.8 Scope and limitations

**Chapter 2 Background of the Study**

2.1 The Institute

2.2 Library and Information Centre

 2.3 Information Resources

2.4 Information Services

**Chapter 3 Evaluation of Existing System**

3.1 Introduction

3.2 Evaluation of the System

**Chapter 4 Recommendations and Concluding Remarks 20**

* 1. Introduction
	2. Recommendations to Improve the Existing Repository
	3. Re-designed Digital Repository

 4.4 Concluding Remarks

 **References**

**Chapter 1**

**INTRODUCTION**

Institutional Repository (IR) also known as digital repository or digital library is a solution for gathering, preserving and disseminating intellectual output of an educational and research organization, and Online Public Access Catalogue (OPAC) is a tool for preserving and retrieving the bibliographic elements of its collection of information resources. The availability of user-friendly Open-Source Software applications, i.e. the applications with their source codes open for inspection, modification, and improvement by adding features to them, and increase in the I.T. literate information user community have made IRs and OPACs a common place in educational institutions across the world.

* 1. **Institutional Repository: Concept, Definition and Content Types**

Lynch (2003) defined an institutional repository as 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”. According to Swan, they are “digital collections of the outputs created within a university or research institution”. The IRs are built on a set of global technical standards which make them ‘interoperable’. These standards are known as Open Access Initiative-Protocol for Metadata Harvesting or OAI-PMH, in short. Earlier, the IRs were considered as online channels for the dissemination of ‘grey literature’, the unpublished theses, dissertations, term papers, presentations etc. of an organization which are typically not available or accessible for the scholarly community outside the organization. However, the recent trend is to promote IRs as and other open access repositories as primary channels of scholarly communication, and thus reduce the monopoly of commercial journal publishers in the scholarly world.

Any scholarly piece of information, generated at an Institutional level is qualified to incorporate in the Institutional repository. In general, IRs have the following types of resources.

1. Theses
2. Dissertations
3. Conference papers
4. Conference Presentations
5. Teaching notes
6. Class presentations
7. Peer-reviewed journal articles or pre-prints
8. Research datasets
9. Books and Monographs

**1.3 Advantages of Institutional Repository**

The major advantages and benefits of establishing an Institutional Repository are the following.

1. Dissemination of scholarly output of the organization to a wider audience
2. Increase in the citation to the research works
3. Showcasing of the scholarly performance of the organization
4. Long-term preservation of the scholarly content
5. Opening of opportunities for collaborative and inter-disciplinary research

1.2 Online Public Access Catalogue: Concept and Definition

The Online Public Access Catalogue or OPAC in abbreviation, is the computerized version of traditional library catalogue. Typically, OPAC functions as a subsystem or module of a software package known as Integrated Library Management System (ILMS) or Integrated Library System (ILS) which is used for managing the entire library operations and services. Chow & [[Anthony S. Chow, Timothy Bucknall, in [Library Technology and User Services](https://www.sciencedirect.com/book/9781843346388/library-technology-and-user-services), 2012]]

Advantages of the Online Public Access Catalog

does not need to be sorted statically; the user can choose author, title, keyword, or systematic order dynamically. 2. Most online catalogs allow searching for any word in a title or other field, thereby enhancing the ways to search and find a record. 3. Many online catalogs allow links between several variants of an author's name. For example, if an author's name "John Adams Smith" is keyed in differently also as "John A. Smith" and "J.A. Smith" in a catalog, there would be links between the three names thereby enhancing find ability. 4. The elimination of paper cards has made the information more accessible to many people with disabilities, such as the visually impaired, wheelchair users, and those who suffer from allergies that are paper- or building-related. 5. Physical storage space is greatly reduced. 6. Updates are easier and more efficient

**Current international and national status of Institutional Repositories and Online Public Access Catalogues**

There has been a proliferation in the number of IRs and OPACs both internationally and nationally in the last few years with the emergence of Open Source Software solutions in the field. According to the Ranking Web of Repositories (2016), totally there are 2297 no. of IRs all over the world and in India there are 43 repositories. These repositories differ greatly in content, size, scope and objectives and majority of the IRs operating in India are open source based (Gohain 2011). Two active national repositories in India at present are Sodhganaga, a repository of Indian doctoral theses and Sodhagangotri , a repository of Indian doctoral research in progress, both developed by the INFLIBNET, a constituent of UGC, Govt. of India. The National Digital Library is a new repository from India developed at the IIT, Kharagpur as a part of the National Mission on Education through Information and Communication Technology of the Ministry of Human Resource Development, Govt. of India.

**1.4 Statement of the problem**

All India Institute of Speech and Hearing (AIISH), the premier institution in the country in the field of communication disorders established in the year 1965. The Institute functions under the Ministry of Health and Family Welfare, Government of India. The Library and Information Centre of the All-India Institute of Speech and Hearing, Mysore, has been in existence since the establishment of the Institute on 9th August 1965. It is rich in print and electronic information resources and provides a number of traditional and electronic information services. The library is running an Institutional repository of theses and dissertations. Also, all its operations have been automated using an Integrated Library management System with an OPAC module. Both the applications were developed by third party using proprietary software tools. Presently, the repository contains more than 1700 records consisting of postgraduate and doctoral dissertations, and the funded research project reports. The OPAC contains the bibliographic records of 20,000 plus print collection of the Institute library. Both the IR and OPAC are being maintained on payment basis by the third parties who developed them. Our working experience with the applications and the feedback we have been receiving from our user community strongly suggest a thorough modification in the functionalities and upgradation of both the digital repository and OPAC. However, upgradation and modification processes require additional expenses from the part of the Institute as the applications are running on proprietary software developed by third parties and getting it done by making payment is not feasible considering the fact that the upgradation may be needed frequently. On the other hand, there are plenty of free open source applications for building digital repository and OPAC with advanced features and functionalities. Upgradation is also free on these software. Hence, this project proposes to design and develop an open source platform for the digital repository and OPAC and migration of the existing digital repository collection and the OPAC bibliographic records to the new system.

**1.5 Objectives of the Study**

The objectives of the project are the following.

1. To systematically evaluate and summarize the drawbacks of the existing IR and OPAC of AIISH.
2. To identify suitable open source software alternatives for the existing IR and OPAC.
3. To design and develop an open source platform and migration of the resources
4. To expand and enrich the content base of the repository by additionally incorporating new material types
5. To facilitate content based searching of digital repository and discoverability on search engines.
6. To generate a new comprehensive set of metadata for the IR and OPAC as per the international standards
7. To facilitate interoperability for the IR with the related institutional and national repositories across the world.
8. **Importance of the Present Project**

A review of the current international and national status of the subject reveals the importance of developing Institutional Repositories. Also, majority of the repositories that have been developed in India are on open source platform and they have a lot of features and functionalities such as content based searching, e-mail alerts, interoperability with other IRs, web search engine discoverability, etc. which AIISH digital repository does not have. Hence, it is imperative to develop a repository for the Institute on open source platform for the benefit of its user community.

**4.0 Work Plan**

**4.1 Materials and Methods**

1. **Analysis of the existing Digital Repository and OPAC**

The present system of digital repository and OPAC will be analyzed to identify its strength and weakness by the following methods.

1. Conducting an evaluation survey among the end users on their usability.
2. Testing and evaluation of the system against a set of functional requirements as per the Digital Repository Policies and Functional Requirements Specifications of the National Library of Medicine (NLM), USA. (National Library of Medicine 2007).
3. Comparing the performance with a representative samples of other institutional repositories and OPACs in India and abroad.
4. **Selection of Open Source Software**

The available Open Source Softwarefor building digital repository and OPAC will be identified by referring Source Forge, an online directory of open source software (Find, create and publish, n.d.) and also by conducting a literature survey in the field of digital information management. After preliminary evaluation of the identified software applications, three will be shortlisted as candidate software. In-depth testing and hands-on analysis will be carried out with the candidate software and one among them will be chosen for the development of the system. The other associated open source software applications needed for developing the system namely, MySQL, PHP, Apache and Tomcat will also be downloaded and installed.

1. **System Design and Development**

Based on the analysis, a new system will be designed, taking into consideration of the limitations of the existing system and the need for incorporating new resources such as photographic images, videos, presentation slides, question papers and syllabi with additional features and functionalities.

1. **Customization and Building up of the System**

The software applications selected will be integrated, customized, modified and additional features will be incorporated by writing necessary codes/ programmes. MySQL based databases will be set up and necessary configurations will be made on other software applications. Also, web based interfaces will be designed and developed for both the repository and OPAC.

1. **Rectification of the Digital Repository Content**

The digitized materials (non-born digital) of the existing repository (approximately 500 reports) will be thoroughly checked for errors such as spelling mistakes and missing of contents. This will be done manually as well as by using ABBYY FineReader, the optical character recognition software. The errors identified will be rectified with the help of a Speech and Hearing Professional. Also, the entire contents of the digital resources in the repository will be watermarked as a part of Intellectual Property Rights management.

1. **Rectification of the OPAC Records**

In order to ensure accuracy of information and improve the OPAC based search results, the existing bibliographic records of the print collection (more than 20,000) will be completely replaced by importing the corresponding MAchine Readable Catalogue (MARC) records from the official website of the Library of Congress, USA ( Library of Congress, n.d.) available for free downloading as a part of its Cataloging in Publication programme. The records which are not available with the Library of Congress will be created manually.

1. **Metadata Creation**

Metadata are the elements for describing information resources in a collection. In order to index and describe the resources included in both the Digital Repository and OPAC, appropriate metadata will be created from the Dublin Core Metadata Set, an international standard for information resource description developed by Dublin Core Metadata Initiative (DCMI), a non-profit organization and ratified as ISO 15836:2009, ANSI/NISO Z39.85-2007, and IETF RFC 5013 (DCMI n.d.). Metadata such as contributor, coverage, description, year of publication, format, publisher, subject, title, ISBN, accession number and location will be considered for inclusion.

1. **Uploading of the Materials**

All the materials will be uploaded to the new system developed by assigning suitable metadata elements selected. Wherever possible bulk uploading will be done by writing necessary scripts in PHP language.

1. **Trial Running, Error Rectification and Finalization of the System**

Trial running will be carried out after development of the system. Then, errors, if any will be rectified and the system will be finalized.

**Chapter 2**

**Analysis of the existing Digital Repository and OPAC**

 Communication disorders are difficulties involving speech, language and or hearing, which have been accepted as a major public health issue as they compromise early childhood development, restrict vocational attainment and attenuate the economic well-being of the society.

 The All India Institute of Speech and Hearing (AIISH), Mysore, is the foremost organization in the country, mandated to help and support people with communication disorders. Founded in 1965 as an autonomous institute under the Ministry of Health and Family Welfare, Government of India, AIISH has been successfully carrying out its activities for the last 46 years. In pursuit of its mission of ensuring effective communication by one and all, the institute (a) generates and prepares knowledgeable and skilled practitioners in Speech - Language Pathology and Audiology and allied fields, who can provide state-of-the art, client-centered service in multiple settings,(b) conducts both basic and applied research in areas related to communication disorders and disseminates the findings among the local, national and international audiences, (c) provides affordable, accessible and high-quality speech, language, and hearing clinical services to the needy ones, and (d) develops, implements and evaluates education and outreach programs to address the prevention, intervention and control of communication disorders.

 Started as an institute to train the postgraduate students in Speech and Hearing, today AIISH has grown into the largest institute in the country offering various diploma, graduate, postgraduate, postgraduate diploma, doctoral and post doctoral programs in the field, in addition to rendering clinical services to thousands of persons with communication disorders from all over the country and conducting multi-faceted research to find solutions to the challenges of communication disorders. All these activities are spearheaded by the experts working in 11 departments of the I n s t i t u t e , n a m e l y A u d i o l o g y, C e n t r e f o r Rehabilitation and Education through Distance Mode, Clinical Psychology, Clinical Services, Electronics, Otorhinolaryngology, Prevention of Communication Disorders, Special Education, Speech-Language Pathology, Speech-Language Sciences, and Material Development.

The institute has been recognised as a Centre of Excellence in the area of deafness by World Health Organization, as a Centre of Advanced Research by University Grants Commission, as a Science & Technology Institute by Department of Science and Technology, Government of India and as the nodal centre for implementation of prevention and control of deafness by Ministry of Health & Family Welfare, Government of India.

**2.2 Library and Information Centre**

 The Library and Information Centre of the All India Institute of Speech and Hearing, Mysore, has been in existence since the establishment of the Institute on 9th August 1965. The centre is housed in a four-storeyed building with an area of 24000 square metre. The ground floor houses the Book Section, Back Volumes of Journals, and a browsing center, the first floor houses a section for display of new books, Current Journals and Reference books, thesis, dissertations, independent projects and research projects, the second floor houses the digital library workstation and two spacious reading halls and the 3rd floor houses the computer browsing center. The building and the furniture are specially designed to meet functional, as well as the aesthetic requirements. The building is surrounded by beautiful lush green lawn and well-maintained rose garden and ornamental plants.

The library is well equipped with conventional resources, as well as technology based information services. The AIISH library and information center is a specialist documentation center for the rehabilitation professionals in general and professionals in the area of communication disorders, in particular. The center has received more than eleven hundred books and journals on speech and hearing on gratis from Wendell Johnson library, Iowa, USA. The Library also got support through Vocational Rehabilitation Administration (VRA) project started at the institute in 1967 The library has been using the Book Magic an Integrated Library Management software package with all the modules of library housekeeping operations. All the holdings of the library have automated access and users can search the library database by using OPAC.

**2.3 Information Resources**

The Library and Information Centre (LIC) of the institute houses a rich collection of information resources pertaining to the field of Speech and Hearing and allied areas of Clinical Psychology, Otorhinolaryngology and Special Education. It has both print and electronic formats of documents and caters to the needs of students, research scholars, teachers and practitioners, not only of the institute but also across the country. Library and Information Centre web portal acts as a gateway to a diverse kinds of electronic resources pertaining to the field like E-journals, E-books and online databases. In addition, the LIC has been maintaining an online digital repository of research works carried out in the institute and making them available free of cost to the professional community worldwide.

1. **Print Books: The library procures all the books published by publishers all over the world pertaining to the field of Speech and Hearing. Presently it has more than 15,000 books exclusively on speech and hearing in addition to the 3000 books on allied areas like clinical Psychology , Special education and Otorhinolaryngology.**
2. **Print Journals: It subscribes to 64 print journals pertaining to Speech language pathology, Speech Language Sciences, Speech Communications, Speech processing, Audiology, Special Education, Clinical Psychology and Otorhinolaryngology.**
3. Bound Volumes of Periodicals: The library has a good collection of bound volumes of periodicals pertaining to **speech and hearing in addition to that of allied areas like clinical Psychology , Special education and Otorhinolaryngology.**
4. **Theses and Dissertations : These constitute another valuable resource of library and information centre and include PhD theses,master’s dissertations and other reports of various project works sponsored by the national and international agencies like Department of Science & Technology,Government of India, Rehabilitation Council of India and World Health Organization.**
5. **E-Books: The library procured 138 E-books pertaining to the filed of Speech and Hearing and allied areas published by 4 publishers namely Taylor and Francis, Oxford University Press, Wiley Interscience and Elsevier.**
6. E-Journals: The **library** has subscribed 63 E-journals. Among them, 48 are available in print format also.. In addition to the subscribed E-journals, **library** is getting access to more than 2500 biomedical E-journals as it is a part of Electronic Resources in Medicine consortium, launched by Director General of Health Services, Ministry of Health and family Welfare, Government of India.
7. **Bibliographic databases: The library subscribes two bibliographic databases namely i)** ComDisDome an indexing and abstracting tool covering thecommunications disorders literature, with focus on speech-language pathology and audiology. Quickly identify and connect to reliable information from multiple sources , including journal articles and books, along with access to profiles of researchers working in this area and (ii) Linguistics and Language Behavior Abstracts (LLBA).This database abstracts and indexes the international literature in linguistics and related disciplines in the language sciences. The database covers all aspects of the study of language including phonetics, phonology, morphology, syntax and semantics. Documents indexed include journal articles, book reviews, books, book chapters, dissertations and working papers.
8. CD/DVD-ROMs: The library has a good collection of both Comapct Discs and Digital Video Discs. Most of these resources are part of the text book collection in the library.
9. Microfilms: The library has a rare collection of Speech and hearing related microfilms. The total numbers of microfilms are more than 200.
10. Digital Repository: The Library is maintaining an online digital repository of more than 1000 research works carried out in the institute and making them available free of cost to the professional community worldwide.

**2.3 Information Services**

**1.Traditional Information Services**

 The major traditional information services provided by the LIC are the following:

1. Reference Service

 The Library and Information professionals provided assistance to hundreds of users in

 matters like locating required information, evaluating the information resources and

 usage of resources. They answer various reference queries in person, via e-mail and

 over phone.

1. Book Lending Service

 The book lending is the most sought after service of LIC. On an average, 200 books

 were issued daily to the users.

1. Reprographic Service

 This is one of the most heavily used library service. Thousands of library materials

 are being reproduced daily on payment basis under this service. Only the library

 materials are permitted to copy and reproducing complete document was prohibited as

 it violated Copy right Act.

**2.Electronic Information Services**

 LIC provide two categories of electronic information services to its users. These are: (1) Internet Service (II) Web portal based services.

* 1. Internet Service

 LIC strengthened the provision of Internet service to the member community by adding more number of computers with Internet access. Totally there are 44 computers with Internet access in the LIC.

* 1. Web portal based Information Services

Library and Information Centre provided a variety of information services through its web portal available at **www.aiish.ac.in**. Some of them are listed below:

1. Access to Digital Research Repository

 The institutional digital repository developed and maintained by the LIC continued to raise the visibility of the research work done at AIISH to scholars around the world. There are nearly one lakh pages of research reports in the repository.

1. Web Online Public Access Catalogue Service

 The bibliographic details of over 17,000 books and bound volumes of periodicals were made available over the web. Thus the users all over the world can search the information resources of the institute using search terms like author, title, publisher, and keywords.

1. Access to E-Journals, E-books and Online Databases

All the subscribed E-journals, E-books and Online databases have hyperlinked

through the web portal. In addition to the subscribed resources, the portal also provides access to high quality information resources pertaining to Communication Disorders which are available free of cost over the web like **‘REHABDATA’** a leading literature database on disability and rehabilitation published by National Rehabilitation Information Center, USA.

**Chapter 3**

**Evaluation of the Existing SystemS**

**3.1 Introduction**

 The existing digital repository of the institute have been launched in 2008 with an aim of long-term storage and archiving of our valuable research outputs in the field of speech and hearing. It also aims at easy dissemination of these resources to the rest of the world.

 The repository is made up of Equest , a customized software solution for digitized contents of the libraries (Thesis, Projects, Research Papers, and Publications). It provides search and access to digitized projects, such as thesis, research paper, publications etc. Can host and manage e-contents on Internet along with online payment facility. The e-contents files could be in various formats like word, text, HTML, PDF etc. Can be implemented on LAN / Intranet, Users can submit their thesis, Research Papers, Projects online. The user can save an article to his or her e-library and can also email the articles to others. PDF files can be downloaded. Search facility includes Basic & Advanced. Basic includes 3 parameters and Advanced have 12 parameters. Users can search on the entire database of the software, Searches can be performed on 16 different search fields based on time frame, subject, departments, publication type, and publication dates. Equest is complied or compliant? Or Equest complies with international metadata and interoperability standards: MARC-21, MARC-XML, MODS, DUBLIN CORE.

 The repository is named as ‘Digital Library’ and is available on the Library and Information Centre portal at [www.aiish.ac.in](http://www.aiish.ac.in)

**3.2 Evaluation of the System**

An evaluation of the existing system identified the following:

**1. Platform**

The present repository is built using proprietary digital library software titled ‘Equest’ designed by an Indian software company, Total IT Solutions Private Limited, Pune. Equest is a customized software solution for digitized contents of the libraries .This software helps in managing e-contents. Publishers / Institutes can also use this software to host & manage their e-contents on Internet.

1. **Contents**

The major contents available on the repository are the following:

1. PhD theses: The repository houses 56 PhD theses pertaining to the field of Speech and Hearing done by the staff and research scholars of the institute under the able guidance of the experienced faculty members of the institute.
2. Dissertations: This is the second category of digitised resources in the repository. These are the research works carried out by the Post Graduate students of various courses in Speech and Hearing and allied sciences. The total numbers of digitised dissertations in the collection are 767 records.
3. Department Projects: These are the funded research project reports carried out by the faculty and staff of the institute. The funded organizations are either the institute itself or external national and international organizations like Rehabilitation Council of India, Department of Science & Technology, Government of India and World Health Organization.
4. Independent Projects: Independent projects are the research works done by the first year Post Graduate students of previous batches.

No other types of materials have been included in the repository.

1. **Number of Digital Materials**

The repository contains only 1300 reports under the categories of PhD theses, dissertations, department projects and Independent Projects.

1. **Content Recruitment Methods**

The only content recruitment method followed by the repository is the uploading of materials by library staff.

1. **Search Features**
* The repository offers both simple and advanced search facilities. The Boolean operators AND,NOT,OR based searches. It is possible to browse the repository by author,title, [Author](http://aiishdigilib.in:8080/digitallibrary/HomePublication.do), [Guide](http://aiishdigilib.in:8080/digitallibrary/HomeGuide.do), [Subject](http://aiishdigilib.in:8080/digitallibrary/HomeFaculty.do), [Resource](http://aiishdigilib.in:8080/digitallibrary/HomeResourceWise.do) type and [Year](http://aiishdigilib.in:8080/digitallibrary/HomeYearDistinct.do).
1. **Conversion of Print Matter into Digital Format**

The print matter to be incorporated into the repository are converted to digital format by the library staff using a high capacity Epson Scanner.

**7. File Formats**

The repository files are in portable document format. All the materials to be deposited in the repository are submitted either directly as PDF or other file formats like MS WORD which will be converted to PDF by library staff.

**8.Access Details**

 Anybody can access the contents of the repository from anywhere in the world through the repository interface at [www.aiish.ac.in](http://www.aiish.ac.in). No restriction has been made in this regard. Full-text access is provided to all the resources in the collection.

**9.Submission of materials to the repository**

Submission materials to the repository have been done by the library staff only. There is no provision for the faculty members and researchers to submit their work by themselves on to the repository.

**10.Policy Formulation Regarding Digital repository**

 No policy has been formulated with regard to the institutional repository so far.

**11.User friendliness**

The existing system lacks user friendliness as there is no help link or user manual.

**Chapter 4**

**Recommendations and Concluding Remarks**

**Introduction**

In this chapter recommendations have been made to improve the existing digital repository of the institute. The chapter also proposes a new redesigned digital repository on Eprints, a well known digital library software used worldwide.

* + - 1. **Recommendations to Improve the Existing Repository**

Based upon the evaluation of the existing system, the following recommendations have been made to enhance the usage of vital speech and hearing digital repository of the institute.

**1. Policy Formulation**

The institute has to draft policies have to be drafted regarding the following areas of repository.

* Defining collections
* Intellectual property
* Determination of what is acceptable content
* Determining who is authorized to make contributions to the repository
* Acceptable file formats
* Metadata formats and authorized metadata creators

**2.** **Content-recruitment methods**

The availability of the repository may be publicised in the campus through notice boards, official communications, institute newsletters and other means of communications. The best content recruitment methods that suit the activities of the institute may be identified. Harvesting of publicly available materials may also be considered. While the faculty members and staff may be permitted to upload their research works directly, the submission of post graduate and PhD theses may be done through the faculty supervisors.

3. **Requirements Related to Submission of materials**

The following requirements address the entire process of authorized users submitting learning objects into the repository. The submission process should involve logging in to the repository with an authorized username and password, creating metadata records for items submitted, uploading the files associated with a learning object into the repository, and workflow related considerations.

**4. Inclusion of Varied Types of Contents**

The repository may incorporate various other materials like class lecture and guest lectures in audio and video files, study materials, question papers, syllabus and other learning aids.

**5. Number of Materials in the Repository**

Presently the repository has only 1200 files of digital materials in the repository. The number may be enhanced by incorporating all the educational and research materials brought by the institute in digital format.

**6. Acceptable File Formats**

Digital materials in the following formats may be accepted for inclusion in the repository.

a. Portable Document Format(PDF)

b. Joint Photographic Experts Group (JPEG)

c. Tagged Image File Format (TIFF)

d. Graphics Interchange Format **(**GIF**)**

e. eXtensible Mark-up Language(XML)

f. Microsoft Word

g. Microsoft Excel

h. Rich text

i. Microsoft PowerPoint

j. Postscript

k. Moving Picture Experts Group **(**MPEG**)**

l. Plain text ANSI X3.4/ECMA–6/US-ASCII (7-bit)

m. Plain text UTF–8 (Unicode)

n. Plain text ISO 8859–x (8-bit)

o. Plain text (all other encodings, including, but not limited to, ISO 646, national variants)

p**.** Portable Network Graphics (PNG)

q. TeX

**7. Requirements Related to End User Access to Learning Objects**

The following requirements address end user interaction with the repository. All users should be able to search and browse the repository and at the least see the metadata records associated with the learning objects, even if not all users will have access to the files associated with the learning objects.

• Search full-text and metadata elements

• Browse by various access points or categories

• View simple and complete metadata records

• Access learning objects as allowed by user account authorization

**8.Encouragement and advocacy**

The usage of the repository may be encouraged by organising various user awareness and training programs.

**9. Enhanced User friendliness**

There should be help functionality in the repository directing users on how to use and navigate through the repository.

**10. Search facility**

There should be a simple search interface which has a link to the advanced search facility. Presently both simple and advanced search facilities are combined.

**11. Repository Administration Requirements**

 The following requirements address various administrative tasks and responsibilities for managing the repository.

• Manage user accounts

• Set authorizations and permissions

• Implement organization structure for learning objects

• Customize submission workflow

• Implement metadata scheme

• Customize metadata input

• Administer and manage items in repository

**B. Proposed Redesigned Repository**

**1.Software Selection**

An evaluation of available open source software platforms were conducted in order to make an informed decision on the most appropriate platform which will fulfil in the needs of the institute. The following software were reviewed:

1. DSpace
2. ePrints
3. Fedora
4. Greenstone

It is decided to use Eprints 3.3.6, a free and open source software package for building open access repositories that are compliant with the Open Archives Initiative Protocol for Metadata Harvesting. It shares many of the features commonly seen in Document Management systems, but is primarily used for institutional repositories and scientific journals. EPrints has been developed at the University Of Southampton School Of Electronics and Computer Science and released under a GPL license.

 As the first professional software platform for building high quality OAI-compliant repositories, EPrints is already established as the easiest and fastest way to set up repositories of open access research literature, scientific data, theses, reports and multimedia. EPrints is a major leap forward in functionality, giving even more control and flexibility to repository managers, depositors, researchers and technical administrators.

The following features of Eprints3 have been considered for selecting the software as the platform for the proposed repository.

### a.Importing and Exporting Data

 There are numerous format options for exporting search results including Dublin Core and METS as well as bibliographic software formats like Reference Manager. These are implemented using plug-ins described as 'kind of cool', written by Southampton, but to which other users can also contribute.Another 'cool' feature is latitude and longitude fields which can be used to export data to Google Maps and look like fun to play with. In the context of scholarly publication, the potential uses still need thinking through. The demonstration showed locations on a Google Map using co-ordinates added into some sample records, but this was hardly a serious application. There may be scope for using the feature to show the location of field research sites, assuming the depositors/mediators are prepared to provide the necessary metadata. It is an indicator though of the way in which EP3 is evolving to provide hooks into other services and to provide new opportunities to work with the material deposited. In summary, there have been some interesting improvements to the search and browse interfaces, but they are, as yet, of unproven usefulness.

### b. Deposit and Workflow

There have been major changes to the deposit process, in order to make it more user-friendly and to encourage further take-up of self-deposition. This is one of the key changes in EP3 and the one which could potentially have the greatest impact on those institutions which are using a mediated deposit model. When implemented in conjunction with the autocompletion feature, it provides an easier interface for self-depositors.The first major change is the introduction of a set of tabs for the various deposit stages:

Type -> Upload -> Details -> Subjects -> Deposit

 These feel similar to the 'sausage bar' used in DSpace's deposit process . The tabs are effective because they make it clear from the outset what stages are involved. In principle they also show depositors where they are in the process, although the stages do not necessarily have to be completed in the order given, and repositories can customise the sequence. There is error-checking for obligatory data before final deposition.

**c. Autocompletion**

There was much play made of various fields having autocompletion. The default ones are journal, author and ISSN. Autocompletion uses JavaScript to monitor what is being typed and queries the relevant EPrints MySQL tables. This means that as more data is added to the archive, the more useful autocompletion becomes. In theory, the autocompletion features could be made to query an external database (such as Zetoc or SHERPA/RoMEO), but network connections are likely to be too slow for this to be effective. One exception might be using LDAP on a local network to validate 'creators'. The autocompletion for journals in the demonstration archive was linked to Southampton's version of the RoMEO database. Matched journals were displayed along with some RoMEO data and Southampton's version of the RoMEO colour codes. Another approach could be to create pre-populated local databases of researchers and staff, journals, etc., from such sources as the institution's publications database, so that they can then be used for look-up and validation. In the case of journals, there is an opportunity for external databases to offer a facility for checking titles in bulk and returning appropriate data for storage locally. This could be done through an API and/or something new such as a Web form where a list of titles could be copy-and pasted for processing. Such a facility would also have other uses. The autocompletion feature, for authors' names in particular, raises an interesting authority issue. Should the repository use the authority name for an author or his or her name as cited in the published paper or research output? Perhaps there is scope for a separate author authority field? Autocompletion can also be combined with conditional workflows to provide a customised self-deposit process.To select a subject or academic unit, the user is no longer presented with one massive hierarchical list. Instead, only the top levels in the hierarchy are displayed, which can be expanded/contracted to show the lower/higher levels. This approach is similar to the way folders can be expanded and contracted in Windows Explorer. It can also be applied to other EPrints pages where little-used fields can be collapsed out of the way until needed. What is more, these views can be made conditional. So, for instance, for a depositor from the History Department, the history subject classes could be automatically expanded to their full depth while keeping the other subjects in a collapsed state.

 Poor usability is one of the barriers that deters authors from self-archiving their publications. The suite of new deposit, workflow and autocompletion features in EPrints 3 have gone a long way in making the deposit process much more user-friendly and intuitive. It will be interesting to see over the next year if repository administrators can exploit these new features to encourage more self-deposit into their repositories.

**2.Operating System**

 Originally EPrints is developed on Redhat Linux but it is used on any number of Linux distributions, and other UNIX-like systems including OS-X. It also is supported by Microsoft’s Windows operating system namely , Vista XP and Windows.The preferred operating system for the proposed repository is Ubuntu 10.10.

**3.Metadata Standards**

Metadata is a core issue for the creation of repositories. Different institutional repositories have chosen and use different metadata models, elements and values for describing the range of digital objects they store. Metadata creation is an important factor to be considered in the preservation of material within a repository. Hockx-Yu emphasizes that "much could be done to consider digital preservation from the outset, to involve the authors in contributing preservation metadata during the creation and ingest process and to embed digital preservation into the repositories work flow, which will ease the later preservation tasks"

The re-designed institutional repository will require mandatory self-archiving by faculty and research scolars. Thus, it will include work from many disciplines, possibly in many formats.Faculty will be the primary creators of metadata in the repository, and will need training in how to create this metadata. The repository will include pre-prints and post-prints.

DublinCore open source metadata software has been identified as the metadata standard for the digital repository, and will be customized according to the needs of the various focus areas.

**4.Digitisation Standards**

Digitisation Standards can be identified by constituting a Digitisation Workgroup and will have to be adapted according to the needs of the various focus areas within Eprints.

**5.Digital Objects**

The following categories of digital objects have been identified to be included in the repository.

**i) Events**

 Exhibition

 Conferences

 Open days

 Workshops

**ii) Images**

Maps

Paintings

Drawings

Plans

Photos

**iii) Research papers/reports**

PhD theses

Journal articles

Project reports

Dissertations

 iv) **Moving images**

 Animation

 Movies

 Videos

**6.Manpower Requirements**

The following categories of personnel are needed to look after the functioning of the repository.

**i)Eprints System Manager**

The EprintsSystem Manager should have the following skills:

* Technical management skills;
* System monitoring, testing, debugging;
* Develop portions of Eprintsrelated to system administration;
* Monitor and upgrade Eprintsutility programs and middleware;
* Develop approved system enhancements;
* Manage hardware contracts and system administration;
* Java programming;
* Networks;
* Unix/Linux Server

**ii)Eprints User Support Manager**

The EprintsUser Support Manager will be responsible for:

* Client Support Training;
* Coordinate and manage the definition and setup of new Eprints communities;
* Plan and implement usability tests;
* Make recommendations on new functionality for Eprints;
* Chair the EprintsWorkgroup;
* Write and maintain user documentation for the system, help pages.

**iii)Web Manager**

* Apply usability and user interface design knowledge and expertise;
* Integrate Eprintsinto institute web-site

**iv)Metadata Specialist**

* Share knowledge and expertise about Qualified Dublin Core (as implemented by Eprints) and consult with the EprintsUser Support Manager on questions, issues related to the MARC to Dublin Core metadata crosswalk;
* Adaptation of metadata elements in DublinCore registry of Eprints according to Metadata Std document;
* Provide training.

**v)Digitisation Specialist**

* Serve as knowledge expert for digital preservation issues;
* Provide training;
* Share knowledge and expertise on matters of archival selection,preservation and UP records policies;
* Provide advice regarding scanning of paper or microform documents to digital formats and reformatting of documents already in a digital format.

**vi)Information Specialists**

* Provide general information about Eprintsas a service of the library;
* Alert users to the information potential of the repository;
* Assist end-users with searching the repository;
* Answer end-user questions about Eprints;
* Provide information about the possibility of contributing to the repository

**CONCLUDING REMARKS**

Digital repositories require ongoing evaluation to determine their quality and new directions for growth. A well designed system and user interface will allow the user to develop an appropriate internalized model of that system, which in turn facilitates users’ learning of and interaction with it. Further, Users need to be involved in the development of an application since its early stages. If their tasks and characteristics are not studied, the product that will be

developed may be technically sound, may look very pretty, but will probably not be usable and user friendly. Hence an early involvement of users in system design, adoption of a user-centered approach and the identification of users’ needs and preferences via a user and task

analysis may invariably be done.

**REFERENCES**

Lynch, C.A.: Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL Bimonthly Report* (2003) 1-7.

1. Andersson, U. Short version of the Sesam report: philosophy and rules concerning electronic archives and authenticity.In: Proceedings of the DLM-Forum on Electronic Records, Brussels, 18-20 December 1996. INSAR: European Archives News, Supplement II. Luxembourg: Office for Official Publications of the European Communities, pp. 175-189, 1997.
2. Arms, William Y., [Key Concepts in the Architecture of the Digital Library](http://www.dlib.org/dlib/July95/07arms.html). D-Lib Magazine, July 1995
3. Association of Research Libraries. A Strategic Policy Framework for Creating and Preserving Digital Collections.October 23, 1995.Available at [ahds.ac.uk/ manage/ framework.htm](http://archive.ifla.org/II/%20http%3A/ahds.ac.uk/manage/framework.htm)
4. Aschenbrenner, Andreas, Tobias Blanke, David Flanders, Mark Hedges, Ben O'Steen, [The Future of Repositories?](http://www.dlib.org/dlib/november08/aschenbrenner/11aschenbrenner.html): Patterns for (Cross-) : Repository Architectures.D-Lib Magazine,November/December 2008.
5. Bailey, C.W.Definition and Purposes of a Digital Library.October 23, 1995.
[www.ifla.org/documents/libraries/net/arl-dlib.txt](http://archive.ifla.org/documents/libraries/net/arl-dlib.txt)
6. Bauwens, Michel. Scholarly electronic publishing bibliography.Available at www.
[info.lib.uh.edu/sepb/sepb.html](http://info.lib.uh.edu/sepb/sepb.html)
7. Beagrie, N. and Greenstein, D.A strategic policy framework for creating and preserving digital collections: a report to the Digital Archiving Working Group.British Library Research and Innovation report, 107 London: British Library Research and Innovation Centre, July 1998.[ahds.ac.uk/manage/framework.htm](http://ahds.ac.uk/manage/framework.htm)
8. Bearman, David. Virtual Archives.ICA Meeting, September 1996, Beijing. Draft paper.[www.ifla.org/documents/libraries/net/bearman.txt](http://archive.ifla.org/documents/libraries/net/bearman.txt)
9. British Library. The British Library Information Systems Strategy.November 1995.
[www.bl.uk/services/information.html](http://www.bl.uk/services/information.html)
10. Cassel, Lillian, José Borbinha, [Report on the Seventh ACM/IEEE Joint Conference on Digital Libraries (JCDL 2007) - Building and Sustaining the Digital Environment](http://www.dlib.org/dlib/september07/cassel/09cassel.html): June 18-23, 2007, Vancouver, British Columbia, Canada.
11. Childress, Eric, Carol Jean Godby, Jeffrey A. Young [A Repository of Metadata Crosswalks](http://www.dlib.org/dlib/december04/godby/12godby.html) D-Lib Magazine, December 2004.
12. Crow, R.: *The Case for Institutional Repositories: A SPARC Position Paper*. SPARC, Washington, DC, 2002, 37.
13. Correia, A.M.R., Neto, M.D.: The role of eprint archives in the access to, and dissemination of, scientific grey literature: LIZA - a case study by the National Library of Portugal. *Journal of Information Science* **28** (2002) 231-241.
14. Day, M.W. Preservation of electronic information: a bibliography.
[homes.ukoln.ac.uk/~lismd/preservation.html](http://homes.ukoln.ac.uk/~lismd/preservation.html)
15. Dempsey, Lorcan.Network resource discovery: a European library perspective.British Library. Research and Development Department, 1994.
16. Dempsey, Lorcan, Brian Lavoie, Geneva Henry, [A Service Framework for Libraries](http://www.dlib.org/dlib/july06/lavoie/07lavoie.html). D-Lib Magazine, July/August 2006.
17. Duncan, B. R. (1997). “Characteristics of organizational environment and perceived environmental uncertainty.” In Administrative Science Quarterly, 53(4), p. 404-426.
18. Galloway, Edward A., and Gabrielle V. Michalek. The Heinz Electronic Library Interactive Online System (HELIOS): Building a Digital Archive Using Imaging, OCR, and Natural Language Processing Technologies.The Public-Access Computer Systems Review 6, no. 4 (1995).[www.ifla.org/documents/libraries/net/gale1.txt](http://archive.ifla.org/documents/libraries/net/gale1.txt)
19. Gladney, Henry M., [Safeguarding Digital Library Contents and Users](http://www.dlib.org/dlib/july98/gladney/07gladney.html): Interim Retrospect and Prospects. D-Lib Magazine, July/August 1998
20. Olivier, B., T. Roberts, et al. (2005). “The e-Framework for Education and Research: An Overview”, JISC-CETIS (UK) DEST (Australia). [www.e-framework.org](http://www.e-framework.org)
21. Shanhong, T. (2000), “Knowledge management in libraries in the 21st. century”, 66th IFLA Council and General Conference, Jerusalem, Israel Available at: http:// ifla. inist.fr /IV/ifla66/
22. Yuan Li, Meghan Banach. [Institutional Repositories and Digital Preservation: Assessing Current Practices at Research Libraries](http://www.dlib.org/dlib/may11/yuanli/05yuanli.html). D-Lib magazine. May/June 2011
23. [www.eprints.org](http://www.eprints.org)
24. [www.greenstone.org](http://www.greenstone.org)
25. [www.dspace.org](http://www.dspace.org)
26. [www.aiish.ac.in](http://www.aiish.ac.in)