PROJECT PROPOSAL

PART-A

1.0	**Title of the Project:**	**Design, development and validation of open source platform for AIISH digital repository and online public access catalogue**
	Area of Research:	Others
1.1	**Principal Investigator**	Dr.Shijith Kumar C
1.2	**Principal Co-investigators**	Mr. Nanjunda Swamy, M., Mr. Raghavendra, G.N.
1.3	**Collaborating Institution**	Nil
1.4	**Total Grants required**	Rs. 3 lakhs
1.5	**Duration of the Project**	12 months

2.0 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 proposes to design and develop an open source based platform for the Institute IR and Online Public Access Catalogue and migration of the existing collection of resources to the new platform. The new platform will have provision for incorporating a wide range of Institutional resources such as official photographs, videos, teaching learning materials and question papers in addition to the existing postgraduate, doctoral and funded research projects. It is expected to provide advanced research repository service for the user community and facilitate better visibility for the research output of the Institute at global level.

3.0 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, ie. 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.

3.1 Definition of the Problem

All India Institute of Speech and Hearing (AIISH), the premier institution in the country in the field of communication disorders launched its IR in the year 2009 and introduced the full version of OPAC by installing an Integrated Library management System (ILMS) in 2011. 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.

3.2 Objectives of the Project

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 a open source platform and migration of the resources
4. To expand and enrich the content base of the repository by additionally incorporating the following resources.
    a. Official photographs and videos of the Institute
    b. Teaching learning materials including presentations and guest lectures
    c. Question papers
    d. Syllabus
    e. Any other materials produced at the Institute related to education and research
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.

3.3 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.

3.4 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

I. 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.

a. Conducting an evaluation survey among the end users on their usability.
b. 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).

c. Comparing the performance with a representative samples of other institutional repositories and OPACs in India and abroad.

II. System Design

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.

III. Selection of Open Source Software

The available Open Source Software for 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.

IV. 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.

V. 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.

VI. 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.

VII. Structuring of New Resources

The following new resources will be collected from the concerned Departments/ Sections of the Institute and necessary modifications/ editing/ formatting will be carried out to be inducted to the digital repository.

a. Photographs
b. Video materials
c. Teaching/ learning materials
a. Question papers
b. Syllabus

VIII. 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.

IX. 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.

X. Trial Running, Error Rectification and Finalization of the System

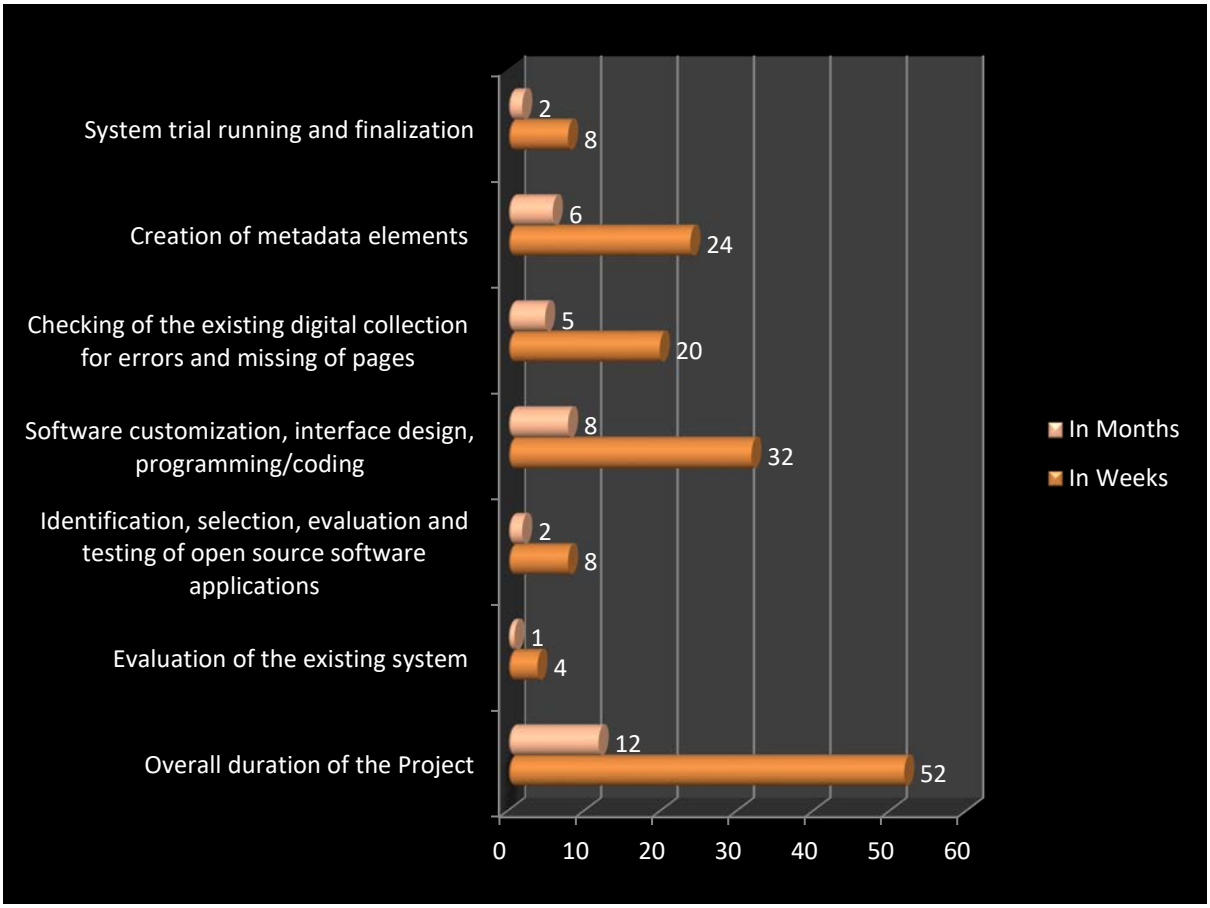
Trial running will be carried out after development of the system. Then, errors, if any will be rectified and the system will be finalized.

4.2 Hardware and Software system requirement

1. Computer with Intel Core-i 3 or later processor, 4 GB RAM and 2 TB storage capacity
2. Open source software for building Institutional Repository and Integrated Library Management System
3. Apache
4. MySQL
5. Ubuntu-Linux Operating System
6. PHP

4.3 Time Schedule

Sl. No.	Activity	Required Time	
		In Weeks	In Months
1.	Overall duration of the Project	52	12
2.	Evaluation of the existing system	4	1
3.	Identification, selection, evaluation and testing of open source software applications	8	2
4.	Software customization, interface design, programming/coding	32	8
5.	Checking of the existing digital collection for errors and missing of pages	20	5
6.	Creation of metadata elements	24	6
7.	System trial running and finalization	8	2



5.0 Budget Summary

SNo.	Item	Expenditure (Rs.)
1	Salary for Research Officer No. of posts: One Qualifications: M.L.I.Sc. Salary: (Rs.25000 x 12 months) Justification: A fulltime research officer with Master's degree in Library Science is required for the project to create and upload metadata for the 20,000 plus records	3.00 lakhs
2.	Consumables	Nil
3	Travel	Nil
4.	Other Costs	Nil
	Total	3.0 lakhs

6.0 Implications of the Study

The present project is expected to have the following implications:

1. Save the annual maintenance charge for the IR
2. Facilitate more visibility for the research output of the AIISH globally
3. Facilitate future upgradation of the AIISH repository to a National Repository on Communication Disorders
4. Facilitate future development of an online union catalogue on communication disorders
5. Help in promoting the development and sharing of digital teaching and learning materials in the area of communication disorders
6. Help other Institutions in planning and implementing new repositories and OPAC
7. Serve as a model for the Institutions who are planning for migrating from proprietary to open source based repository and OPAC.

7.0 Utilization of the Study

The existing institutional repository and OPAC of the Institute in use will be replaced by the new system after getting the necessary approval

8.0 Roles and Responsibilities

SNo.	Investigator	Roles and Responsibilities	Signature
1	Dr. Shijith Kumar C	Overall supervision and coordination, Report writing, Presentation, evaluation of the performance of the existing system, designing of new system, development of workflow, interacting with open source community.	
2.	Mr. Nanjunda Swamy, M	Evaluation of the performance of the existing system, Identification of suitable open source software and testing, System development, Error detection, trial running.	
4.	Mr. Raghavendra G.N	Programming and coding Rectification errors in codes/programs,	
3	Research Officer	Checking all the reports (1700 nos.) and replacing the	

	missing pages by scanning from the print reports. Correcting errors in the content with the help of JRFs. Uploading of reports to the new platform Conversion of the reports from image formats to searchable pdf . Metadata creation for all the reports in the repository. Uploading of reports to the new platform (shared role). Creation of metadata for bibliographic records of the print collection (20,000 plus records)

References

1. Cybermetrics Lab. (2016). Ranking Web of Repositories. Retrieved June 19, 2016 from http://repositories.webometrics.info/en.
2. DCMI Technical Updates. (n.d.). Retrieved August 9, 2016, from http: // dublincore.org/
3. Find, create, and publish open source software for free. (n.d.). Retrieved August 9, 2016, from https://www.sourceforge.net/
4. Gohain Rekha. (2011). Current trend and development of institutional repositories in India. *International Journal of Information Research, 1* ,1.
5. Library of Congress. Cataloging in Publication Program. (n.d.). Retrieved August 9, 2016, from https://www.loc.gov/publish/cip/
6. National Library of Medicine. (2007). Digital Repository Policies and Functional Requirements Specifications. March 16, 2007. Retrieved 9, 2016, from www.nlm.nih.gov/digitalrepository/NLM-DigRep-Requirements-rev032007.pdf