Chapter 1

Introduction
Chapter - 1

Introduction

1.1 Background and Concept

It has been a continual quest of the scholars to identify and have access to the scientific literature in whatever form it is available. Unfortunately, however, the publishers, who solely work for profit making, have largely monopolized the scholarly communication. As it happens they, while accepting the manuscripts for publication get the authorization of the copyright from the author of the work in their favor. It is an irony that manuscripts once published cannot be accessed, neither by the author nor by the institution to which he belongs, without subscribing the journal. The rising cost of the journals and shrinking library budgets result in gradually decreasing access to the journals in the world of scholars. The situation is more alarming in the developing world where libraries have to, somehow, manage to meet both the ends by inter-library resource sharing and other means.

Open access is the direct consequence of the serials crisis emerged during the late 1990s. It was focused on making public-funded research freely available to all without any geographical boundary. The term open access aims at the free availability of the peer-reviewed scholarly journal literature on the Internet so that the users can use the information in any way they like for lawful purposes, transcending the financial, copyright and technical barriers. Open Access is the free availability of research outputs, to anyone, anywhere, without the relentless restrictions on use, normally made obligatory by publishers' copyright agreements. The purpose of open access is to provide barriers free access to the scientific literature. It provides the means to maximize the visibility, ingestion and use of research outputs. The focus of the open access movement is peer-reviewed journal articles, conference papers, datasets of various kinds and any literature of scholarly value. It is assumed that the contents of the documents are available or made available in digital
form that are subsequently made accessible to colleagues, researchers, scientists, students etc. via LAN, WAN, or Internet.

The Budapest Open Access Initiative (BOAI)\(^1\), the pioneer in the field of making scholarly literature openly accessible, suggests the following two methods to achieve the above goals.

1.2 Open Access Journals  
It involves publication of scholarly journals and making them available in digital format for immediate online open access.

1.3 Self-Archiving  
The word archive is synonymous to a repository that holds the digital contents of scholarly information on a server accessible through networks. It involves submission of pre and post peer reviewed version of papers into the archive for Open Access. Such an access is bound to democratize the scientific intellectual capital.

Although there are various types of open access Archives such as Institutional Repositories/Archives, Discipline Repositories/Archives, Author Repositories/Archives etc., the focus of the present study is Institutional repositories. An institutional repository is a set of properly organized and managed collections of digital content produced by faculty members, research scholars, students and other staff of the institution for the purposes of research and teaching. The most important characteristic of institutional repository that distinguishes it from other digital archives is its ‘self-Archiving’. Nowadays, institutional repositories are becoming the best means for capturing and maintaining institutional assets so that it can be openly accessed by the present and future scholars as well.

Institutional repositories hold the promise of being very beneficial to researchers everywhere, particularly to those residing in the developing countries like India, who lack access to academic literature due to its high subscription cost. Also the scholars of developing countries can effortlessly

---

showcase their own research output to the outside world by self archiving it. Although, the major initiatives took place in India for the development of Open Access Repositories such as ‘Shodhganga’, a joint effort of UGC and INFLIBNET to archive Electronic Theses and Dissertations of Indian Universities, yet the growth of open access repositories is in its infancy. It is surprising to note that the number of Academic institutions in India (both in public and private sectors) has grown manifold during the last decade. But the number of Open Access Repositories has unfortunately not grown correspondingly. The two tables\textsuperscript{1} give the present scenario of Academic Institutions vs. OARs, in India as under:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
S.No. & Academic Institutions \tabularnewline
\hline
1 & Universities (Govt., Private and Deemed) \tabularnewline
2 & Colleges \tabularnewline
3 & IITs (Indian Institute of Technology) \tabularnewline
4 & IIMs (Indian Institute of Management) \tabularnewline
5 & IIITs (International Institute of Information Technology) \tabularnewline
6 & NITs (National Institute of Technology) \tabularnewline
\hline
\end{tabular}
\caption{Academic institutions in India}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
S.No. & Open Access Repositories in India & Under ROAR\textsuperscript{2} & Under Open-DOAR \tabularnewline
\hline
1 & Research Institutes & 36 & 24 \tabularnewline
2 & Institutes of Higher Learning (IITs, IIMs etc.) & 33 & 22 \tabularnewline
3 & Universities & 17 & 16 \tabularnewline
4 & Others & 14 & 08 \tabularnewline
5 & Total & 100 & 70 \tabularnewline
\hline
\end{tabular}
\caption{Registered Open Access Repositories in India}
\end{table}

1.4 Statement of Problem

From the foregoing discussion it is quite clear that the number of OARs in India is very small (100 as reported in ROAR and 70 as reported in Open

\textsuperscript{1} Universities, Colleges & Institutes

\textsuperscript{2} ROAR and Open-DOAR are the major Registry services for listing the Open Access Repositories
DOAR) as compared to those available in developed countries (e.g. USA has 574 OARs as reported in ROAR and 440 in Open DOAR). Hence the topic of the present research was selected as “Open Access Repositories in India: Design and Development of a Model”, aimed at studying the present scenario of Open Access Repositories in India, with a view to design and develop a model of OAR by taking in one of the premier Academic institutions of India i.e. Aligarh Muslim University. The problem has been preferred to set up a model of an institutional repository, to identify the problems including efficiencies and inadequacies of developing and launching a repository. The investigator presumes that the present project will assist in establishing OARs in the Academic institutions in general and in Aligarh Muslim University in particular.

1.5 Terms used in the Statement of Problem

1.5.1 Open

According to “Oxford Advanced Learner’s dictionary”, the word ‘open’ is expressed to mean:
1. (Not closed) - allowing things or people to go through
2. (Not hidden)- known to everyone; not kept hidden: an open quarrel, open government. (“Open”, 2000)

1.5.2 Access

According to “Harrods’s Librarians’ Glossary”, the term ‘Access’ is defined:
1. From the point of view of Information Retrieval as
   (i) A device or method whereby a document may be found
   (ii) Permission and opportunity to use a document
   (iii) The approach to any means of storing information, e.g. index, bibliography, catalogue, computer terminal.
2. From the point of view of Archives as

   Availability of government archives to the general public; such documents are subject to restrictions of confidentiality for a specified number of years. Similar restrictions are also sometimes applied to donations or
bequests of other kinds of documents to archive depositories or libraries. Such documents are said to be ‘closed’ until their access date is reached and ‘open’ when the period of restriction has expired. There are other types of access restrictions. Uniform conditions of access are being agreed by the European Union. (Prytherch, 2010)

1.5.3 Repository

The word ‘Repository’ is defined according to “Harrod's Librarian’s Glossary” as:

1. A network-accessible server used as a store for digital content, e.g. in self-archiving or for an E-print archive, and which can disseminate those contents by exposing metadata to harvesters such as the Open Archives Initiative Protocol for Metadata Harvesting. Where the server holds the e-prints for a whole university or similar institution, the phrase institutional repository is frequently used.

2. (Archives) A generic term for that part of an archives service that is concerned with the physical custody and preservation of archival materials. (Prytherch, 2010).

1.5.4 India

According to “The New Encyclopedia Britannica”, the ‘India’ is described as:

Officially Republic of India, Hindi BHARAT or BHARATA VARSHA, that occupies the greater part of South Asia, the seventh largest and second most populous nation in the world. The capital of India is New Delhi. Its area is 1,222,243 square miles (3,165,596 sq km) pop. (2000 EST.) 1,014,004,000. India is a democratic republic made up of 28 states, 6 Union territories and the national capital territory of Delhi. The cultural heritage of India is one of the richest and most ancient in the world. India is a developing country and has a mixed economy in which both the public and private sectors participate. (“India”, 2010)

1.5.5 Design

According to “Oxford Advanced Learner’s dictionary”, the word ‘design’ is expressed to mean:
1. Arrangement: The general arrangement of the different parts of something that is made such as building, book, machine, etc.

2. Drawing/plan/model: The art or process of deciding how something will look, work etc. by drawing plans, making models etc. e.g. the design and development of a model. ("Design", 2000)

1.5.4. Development

The word ‘development’ is defined according to “Oxfoords Advanced Learner’s dictionary” as:

1. Growth: the gradual growth of something so that it becomes more advanced, stronger etc.

2. New product: the process of producing or creating something new or more advanced, a new or advanced product. ("Development", 2000)

1.5.5. Model

According to “Oxfoords Advanced Learner’s dictionary”, the word ‘model’ is expressed to mean:

1. Small copy: a copy of something, usually smaller than the original object; a working model.

2. Description of a system: a simple description of a system used for explaining how something works or calculating what might happen.

3. Example to copy: something such as a system that can be copied by other people, a person or thing that is considered as an excellent example of something (= one that has been specially designed to work well). ("Model", 2000)

1.6 Objectives of the study

In the present scenario Institutional Repositories are becoming a vital part of any institution for preserving its intellectual heritage. The major objective of the study is to describe open access repositories and their development in Indian perspective. This will also entail the study pertaining to open source software packages, relevant for the design and development of OARs. The procedures involved in designing a model of OAR by using a given software will have to be taken into account. The study would be of
Chapter 1 Introduction

immense importance not only for the institutions but also for authors; teachers & students; libraries & librarians.

The objectives of the study are:

- To study the concept and the emergence of OA repositories.
- To make a comprehensive study of Open Access Repositories in India.
- To assess the capabilities and limitations of available open source software packages for the creation of OAR by making a comparative study.
- To work out procedures and parameters for the creation and maintenance of institutional repositories.
- To prepare a model of open access repository for the Faculty of Life Sciences, Aligarh Muslim University.
- To enhance the visibility of the intellectual property and prestige of the institution to the outside world through OAR.

1.7 Scope and Limitations of the study

Libraries are seeking better solutions to cope up with the pressing problems of shrinking library budgets and mounting prices of journals that further lead to declining subscriptions. The situation is so critical that libraries are unable to afford even those journals which are extremely valuable for their users. Also, a large amount of the scholarly literature could not be published because of high publishing cost. The open access movement has brought out new prospects for libraries to deal with the critical situation of unavailability of vital literature to their users due to increasing prices. The open access movement has carried out two innovative ways of barrier free access to the scientific communication. These are Open Access journals or Gold route and Open Access Repositories/Archives or Green route. It depends upon the availability of resources that libraries or institutions give preference to which of the two routes.

Although the subject is well established in the developed world, yet institutions in developing countries are yet to be fully aware of the optimum results of choosing open access options whether in ‘Green form’ or ‘Gold
form' or in both forms. Under the circumstances, a comprehensive study is required on the subject matter that could suggest the scholars and institutions to go for better solutions for enhancing their visibility and prestige to the world wide audiences.

The present work is an effort to make a comprehensive study on Open Access Repositories and also to look into the benefits of preferring open access solutions to resolve the problem of ever-increasing cost of scholarly communication, in which first the author has to pay to publish the article in a journal and then he/she is charged by the publisher to access the same article.

In view of the foregoing discussion, the scope of the study includes the present scenario of OARs in modern libraries followed by a comparative study of open source software packages to bring out capabilities and limitations of each one using the select software, a model of OAR in Aligarh Muslim University, is to be developed. However there are restrictions for the researcher to go beyond a certain limit. In the present case, only one comparatively small faculty of AMU i.e. Faculty of Life Sciences, was selected out of a total number of 12 Faculties. It is a medium sized faculty, with sizable number of publications comparable to those of Arts, Humanities & Social Sciences. However, to develop an OAR of any faculty, the basic data is to be collected from the scholars. As the nature of human beings differs from person-to-person, it is just possible that certain faculty member may co-operate fully while others may not do so during the collection of their research publication. Even in one faculty there may be a large number of research scholars who may be difficult to contact thus depriving some useful publications authored by them. During the collection of data, it was noted that nearly all the publications of the research scholars were in joint authorship with their supervisors/Guides; hence only faculty members were approached to avoid duplication.

1.8 Research Methodology

Research methodology is a manner to systematically resolve the research problem. It describes the various steps that are commonly implemented by a
researcher in studying the research problem along with the logic behind them. It is a principle for solving a research problem, with precise mechanism such as methods of conducting research, tools and techniques used to conduct the study etc. The methodology for the present study is pragmatic in approach that includes the simultaneous use of diverse research techniques to carry out the study. The investigator has made a comprehensive study of Open Access Repositories and gave the detailed account of Open Source Software Packages used to develop Institutional Repositories. The methodology of research is given as under:

1. The concept of OAR was thoroughly discussed in general and in Indian perspective in particular.

2. OARs available in India were surveyed to know their individual structure and functioning with type wise distribution of reported OARs in India as reported in two major registry services i.e. ROAR and Open-DOAR.

3. Documents/Publications were collected personally by approaching each and every Faculty member/Research scholar of the Faculty of Life Sciences, Aligarh Muslim University, Aligarh.

4. The repository so developed was tested by two categories of users, viz. Faculty Members and Research Scholars. A questionnaire was prepared to get the feedback from the Faculty Members as well as Research Scholars belonging to the same faculty.

5. The data collected from the scholars may be in two forms i.e. born digital (pdf) form and in print form. The investigator used the born digital data ‘as it is’, to carry out the study. For archiving the print material, the ‘Digitization’ technique has to be applied and then the files are to be created in PDF (Portable Document Format) format.

6. A comparative study of the available Open Source Software meant for developing OAR was conducted. Therefore, the one found best under the given circumstances was downloaded and installed, followed by metadata encoding and uploading of the collected material.
7. Open Access Initiative Protocol for Metadata Harvesting (OAI-PMH) is a transmission standard for Digital Libraries/Institutional Repositories. As the protocol provides an interoperability framework based on the harvesting of metadata from different repositories, all Open Source Software packages are supposed to be OAI-PMH compliant, a feature that allows the Institutional Repositories/Digital Libraries to provide indexing, search and content description services to their users.

8. The collected data was analyzed and used in the improvement of existing model of OAR.

9. The recommendations for Design and Development of an OAR, especially for an Institutional Repository, is also given.

1.9 Hypotheses

The hypotheses of the present study have not been drafted because of the following arguments:

1. The present study is basically an evaluation and assessment of OSS in general and DSpace in particular.

2. During the course of study, the available Institutional Repositories in Indian Universities till date have also been evaluated. This was followed by the evaluation of 89 Open Access Repositories of Indian origin, registered in ROAR and OpenDOAR.

3. A model IR for the Faculty of Life Sciences, AMU, Aligarh was ultimately designed and developed and put to use by the Faculty Members and Research Scholars of the University.

4. As the data collected for testing the model was confined to one Faculty (i.e. F/O Life Sciences) of AMU, it was not found enough for validation or confirmation of the hypotheses, if formulated.

1.10 Standard Followed

For providing the bibliographical references, American Psychological Association (APA, 2010) (6th ed.) format has been followed. Some examples are given as below:
Chapter 1

Introduction

**Journal Article (print)**


**Journal Article (online)**


**Books (print)**


**Books (online)**


**Web sources**


1.11 Organization of the Thesis

The organization of the study is threefold i.e. Chapterization, Annexure, and Bibliography. The study comprises of nine chapters, each portraying different attributes of the research work under study.

**Chapter 1: Introduction: Methodology**

This chapter introduces the subject followed by the statement of research problem, objectives of the study and the research methodology used to carry out the study.

**Chapter 2: Modern Library: An overview of ICT based services.**

After giving a brief historical development of libraries, the chapter defines and describes digital libraries followed by the procedure involved in
developing a digital library. Issues pertaining to Digital Libraries have been elaborately discussed. Finally, the origin of OAR out of the developments in the field of Digital Libraries has been described.

Chapter 3: Review of Literature

This chapter reviews the literature with a view to study the various trends being followed in the design & development of OAR. The chapter provides complete bibliographical references on various aspects of open access repositories with respect to software packages used for establishing institutional repositories, management of institutional repositories, services, functionality and the creation of OAR model. The relevant material will be collected from research journals, reports, seminar/conference proceedings etc.

Chapter 4: Open Access: An overview

The chapter describes the open access movement, as also the three conferences (Budapest Declaration, Berlin Declaration, Bethesda statement), which laid down the basis for the open access movement and its development. Defining the benefits of open access to scholars, authors, institutions and libraries. Describing the Gold route to achieve open access to scholarly literature followed by Green route of open access. Further, it gives the detailed description of Institutional Repositories, copyright issues, policy mandates and Indian perspective of open access repositories.

Chapter 5: Open source software

This chapter deals with the concept, definition, benefits and licenses of open source software in general. Some select institutional repository software packages have been discussed in the chapter. Apart from this, comparative analysis of four select Software Packages: DSpace, Eprints, Fedora and Greenstone has also been given.

Chapter 6: DSpace, Metadata Encoding and Transmission Standards: A Discussion

The chapter describes the technology pertaining to DSpace such as metadata and its standards that includes Dublin core metadata, METS metadata
along with XML. This chapter further describes the system architecture, functionality, working, and salient features of DSpace software.

Chapter 7: Model IR: Design and Development

This chapter basically deals with the creation of a model of institutional repository for Aligarh Muslim University. It gives a brief description of Aligarh Muslim University and its Faculties including the Faculty of Life Sciences, for which a model has been developed. The chapter provides processes and technical procedures used to create Model IR, Organization, User Registration Process, Self Archiving/Deposition of content, Functionality, Browsing and Searching facilities etc., along with some salient features of the Model IR.

Chapter 8: Evaluation of Model IR, Data Analysis and Interpretation

The Model IR so prepared is put to use before different categories of users in order to collect their feedback. The focus of evaluation appertains to such issues as: general awareness of users regarding Open Access Resources/Repositories, E-Publishing/Copyright in the first part. The second Part covers issues like usage of Model IR, Organization of the Model IR, Publication in IR, Browsing/Searching/Downloading in Model IR, Self Archiving/Deposition of content and usefulness of the Repository. This is followed by Data Analysis, Interpretation and Discussion.

Chapter 9: Findings and Recommendations

This chapter summarizes the whole study, giving the major findings, discussion and recommendations.
Chapter 1

Introduction

References


