ABSTRACT

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.

As per Budapest Open Access Initiative (BOAI), there are two ways of making scholarly literature openly accessible, namely open access journals and
self-archiving. While the open access journals make their digital contents accessible online, self-archiving involves submission of pre and post peer-reviewed version of papers into the archive for open access. The word archive is taken to be synonymous to repository.

An institutional repository which is one of the types of open access archives is defined as a set of organized collection of digital content produced by faculty members, research scholars, students and other staff of the institution for the purpose of research and teaching. In short, an institutional repository serves as one stop source for the total scholarly output of an institution. IRs 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.

**Aims and Objectives of the Study**

In the present scenario Institutional Repositories are becoming a vital part of any institution for preserving its intellectual heritage. 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.


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.

**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.

**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 Open Source Software Systems (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.

Referencing Standard Followed

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

**Journal Article (print)**


**Journal Article (online)**


**Books (print)**


**Books (online)**

Major Findings

The purpose of the present study is to present a model of IR for the faculty of Life Sciences of A.M.U., Aligarh. A protocol was therefore decided to conduct the study. Nineteen open source software packages (taken from ROAR) were studied by the investigator and then D-space software was finally selected for the creation of a model IR. The Procedures and major findings of the study are as follows:

1. D-space was downloaded and installed using windows operating system and other five supporting software packages such as Java JDK, Apache Tomcat, Ant, Maven and PostgreSQL. These five software packages along with Dspace software were downloaded freely from their source sites on the web. Though Linux Operating System is quite stable OS that also supports Dspace, nevertheless, the investigator has selected
Windows 7 OS to build up the repository. To develop a repository a high configuration server or host machine is desirable. The installation process was completed by following the instructions provided on the Dspace source site, i.e. broadly divided into seven steps. After the installation, certain changes were made in the installed package to customize the software as per the requirements of the repository, such as to change Look & Feel, Metadata registry, CSS sheets, image & home page layout, etc. Apart from this, in the technical procedure involved digitization of documents collected in print form.

2. For designing a repository, the scope was limited to the Faculty of Life Sciences, AMU, Aligarh, which is a medium sized faculty, with sizable number of publications comparable to those of Arts, Humanities & Social Sciences. 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. Despite of best efforts, the investigator could only succeed in obtaining the full text publications from a maximum of two Faculty members from each department. Some faculty members directed the investigator to search their articles from Pub-MED Central, Google and Google Scholar. The data thus collected, consists of a total of 521 items including 438 full text Research Publications such as Research Papers, Popular Science Articles, Review Articles, Chapters in edited books and bibliographical details of 83 thesis and dissertations that had been submitted in the departments under study during the period of ten years. A total of 107 articles collected from the Faculty Members in print format were then scanned to convert them in PDF format.

3. The organizational structure of the repository consists of a top level community i.e. Faculty of Life Sciences divided into sub community i.e. the Department of studies; which is further divided into other sub communities i.e. Name of each 'Department of studies', which are
further subdivided into two, Faculty publications and Thesis & Dissertations, which are finally divided into collections.

4. The next step is to upload the publications, which involves seven steps. It is worth mentioning here that there are two methods of deposition of documents in the institutional repository. The one is ‘Through Repository Administrator’ and the other is ‘Self Archiving. Initially the repository was populated by the Repository administrator (investigator), but purpose is to make users/Faculty Members self-reliant in uploading their publications, the provision of Self deposition has also been given to the registered users.

5. The users of the repository have been categorized into three, i.e. General Users, Registered Users and the Repository Administrator. General users may or may not register themselves for accessing and downloading content from the Repository. But the other two types of users are required to get themselves registered through a set protocol of ‘Registration’. Again the registered users may further be divided as the general users who are registered to receive E-mail updates of the recent submissions in the repository and the Faculty Members/Researchers of Aligarh Muslim University who have the privilege to self-archive (or submit) their publications into the Repository along with the tasks. a general user can perform. Besides, the registered users can edit their profile, view their submissions and subscriptions, cancel or save their submissions for later date, get reminder for unfinished submissions.

6. The Repository is capable of Full-Text Searching and Browsing of documents. There are two types of searches, i.e. Simple Search and Advanced Search. The search terms entered by a user in the relevant box are searched by the words given in title, author, keyword and abstract, etc. of each record. The Repository can search full-text i.e. the text entered will be searched in the full-text of all archived publications. The Simple Search can also be restricted to a particular Faculty, Department of Studies and Faculty Members. The Advanced Search option permits a
user to specify the subject areas to be searched, and to combine these searches with the Boolean operators such as AND, OR, NOT, etc. Besides, user can try other searches such as phrase search, combination search using Boolean operators, etc. In Browsing mode a user looks into a particular index, such as the author index, subject/keyword index and navigates around different levels to find the required content. The Browsing preferences are: Browse by Faculty, Department of Studies, Year of Publications, Titles, Faculty Members/Authors/Guides and Keywords.

7. Apart from this, the repository provides certain valuable links and some of the prolific features that add value to the repository such as Help, Copyright, Open Directories and certain Allied links. The users can view the Profile of each Faculty Member, Statistics and Recent Submissions and also may give their Feedback.

8. After a model of institutional repository (Model IR) for the Faculty of Life Sciences of AMU was designed and developed, a questionnaire was prepared to get the feedback from the Faculty Members as well as Research Scholars belonging to the same faculty. The main purpose was to find out the general awareness of both the categories of users with regard to such issues like Open Access Resources/Repository, E-Publishing and Copyright etc. Model IR itself was put to use in order to get the users' reactions regarding its overall organization, Browsing/Searching/Downloading facilities provided therein, Open Access Publishing in Institutional Repository as well as the utility of the process of Self Archiving/Deposition of contents.

9. From the collected data, it was revealed that the users were by and large aware of open access resources. However, not all of them (10% Faculty Members and 4% research scholars) had used these resources before.

10. Awareness regarding the open access publishing was found amongst 50% Faculty Members and 48% Research Scholars only.
11. With regard to using different modes of open access publishing, only 50% Faculty and 48% Research Scholars were reported to be aware of e-publishing only in Open Access Journals. They had no idea whatsoever about publishing their research in institutional repositories, disciplinary repositories and author repositories. Once the importance of publishing in other modes apart from ‘Open Access Journals’ was explained, about 60% users in both categories preferred Institutional Repositories.

12. As to the copyright issues, the awareness was not beyond 45% in any case. The analyzed data revealed that only 20% Faculty Members and 32% Research Scholars had the experience of publishing their papers in e-publishing mode with full knowledge of ‘Author-Publisher Copyright Agreement’.

13. When the model IR was put to use by Faculty Members and Research Scholars from the Faculty of Life Sciences, AMU, it was surprising that none of them had used an institutional repository before. However, once the model IR was used by them followed by explanation of its utility, both the categories of users supported the idea of establishment of a full-fledged Institutional Repository in AMU. However, the users in both the categories were found to be fully satisfied with the organization of various elements in model IR.

14. When asked about their preference of publishing in proposed AMU Institutional Repository, as and when it is launched on permanent basis, 70% of Faculty Members and 100% of Research Scholars answered in affirmative.

15. Regarding the mode of submission of articles in the proposed Institutional Repository of AMU, 40% Faculty Members and 70% Research Scholars preferred to use the method of ‘Self Deposition’, while the remaining users preferred it through ‘Repository Administrator’.
16. Users in both the categories were satisfied with the Browsing and Searching options provided on the home page of Model IR. However, the first preference of search options as given by the Faculty Members was by ‘Keywords’ (100%) followed by ‘Author’ (30%), while the preference of Research Scholars was by ‘Keywords’ (70%) followed by ‘Title’ (34%), Date/Year of publication (12%) and ‘Author’ (34%).

17. As to the active links given in the Model IR, cent percent users were found to be fully satisfied.

18. 100% Faculty Members and 82% Research Scholars found the model IR as user friendly.

19. More than 90% of users agreed that the institutional repository when fully operational will increase the visibility of research and will be able to bring AMU at par with other top universities in the world.

20. 100% Faculty Members and about 50% Research Scholars were in agreement that such a repository will enhance the number of citations vis-a-vis Impact Factor, besides producing better research outputs and improvement in the research techniques employed in various subjects.

Recommendations and Suggestions

During the testing of model IR, some suggestions proposed by the users as well as by the investigator for improvement of the model IR, are summarized as under:

1. Faculty Members and the Research Scholars have found the Model IR of the Faculty of Life Sciences, AMU to be extremely useful for the Institution in general and research community in particular. Therefore, the same should be launched as soon as possible.

2. The proposed IR of AMU should serve as a one stop source for the total academic output of the University.

3. To highlight the Research activity in the University, the proceedings of the Seminars/Conferences/Working Papers of the Workshops conducted in the University should be published online in the proposed IR.
4. Journals published by many departments in the University may be made online using the Knowledge Repository of AMU.

5. Most of the users did not support the categorization of publications on the basis of specialization. They have proposed assignment of more keywords to the indexed documents to facilitate retrieval. They, however, supported categorization, on the basis of type of documents (i.e. books, articles, theses, dissertations, patents, etc.).

6. Under each department, an independent community of Research Scholars/ Research Associates/ Young Scientists may also be added/ created so that their publications, when in single authorship may be listed under that category.

7. Windows 7 platform (OS) was used for developing the Model IR. However, for a full-fledged IR of AMU it is advisable to use Linux Operating System, which is immune to virus attacks.

8. Although all the Faculty members and Research Scholars are prone to ICT, yet users’ awareness programs about ‘General awareness of Institutional Repositories’, ‘Publishing in Institutional Repositories’, ‘Disciplinary Repositories’, including ‘Self Archiving of Publications’ etc., should be regularly conducted.

9. Before embarking upon a full-fledged Knowledge Repository for AMU, well defined policies that include Metadata reuse policy, Full data item policies, Content policies, Submission policies and Preservation policies need to be formulated at the University level. This will also consist of procedures for operation of the repository including methods of submission through Self Archiving and Repository Administrator.

10. The issues of Intellectual Property Right and Copyright should also be covered in the user awareness programs.

11. The staff engaged in designing and developing the Institutional Repository should be technically expert and possess necessary skills such as Digital Collection Management, Open Archive Skills for information System, etc.
Conclusion

In the process of design and development of Open Access Repositories, the objectives of the study as set out in the first chapter have been achieved. Before adopting the set procedure of designing a Model IR for the Faculty of Life Sciences, the study thoroughly covered the concept and the emergence of Open Access Repositories with a specific study of Open Access Repositories available in 16 universities in India so far. A comparative study of the available Open Source Software packages was also made before selecting Dspace Open Source Software for creating a model IR for the Faculty of Life Sciences, AMU, Aligarh. Finally the Model IR was tested by putting it into use by the target users who not only liked the organization of elements on the home page of the Model IR but also the Browsing & Searching options, particularly the process of updating through Emails and Self Archiving/Deposition of contents.

It is suggested that librarians who possess digital collection management and Open Archive Skills for Information System (OAIS) management be recruited. Proper training of Faculty and students to use OAIS, helping them prepare their digital products, involving them in institution-wide policy making and setting repository goals be also provided. It may be concluded that the development of an Institutional Repository System resting on open source software, frameworks and application program interfaces could lower the development cost and time and gives impressive results by addressing the specific needs. In addition, bringing digital assets into a managed repository framework provides new opportunities for digital preservation and is the assurance to resist against technology obsolescence.