ISSN No: - 2456 - 2165

Knowledge Repository: A prototype for Institutional Repositories

Prof. (Dr.) Shabahat Husain
Professor and Chairman, DLIS,
Aligarh Muslim University, Aligarh

Dr. Shazia Khan
Post Doctoral Fellow, DLIS,
Aligarh Muslim University, Aligarh

Abstract:-

A. Purpose :- **The purpose of this paper is to discuss the experiences obtained in developing a proposed model of institutional repository at Aligarh Muslim University, Aligarh and suggestions derived to overcome the challenges.**

B. Design/methodology/approach :- **The basis of this paper is a case study that gives the detailed account of the development of Knowledge Repository, an Institutional Repository of Aligarh Muslim University that includes software installation, customization, workflow in the repository, etc.**

C. Findings :- **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 several 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.**
Research limitations/implications – The development of an Institutional Repository for the Aligarh Muslim University, Aligarh is in its test phase, and, at this moment, only a model Repository by taking the Faculty of Life Sciences as a base has been created.

D. Practical implications :- **The Knowledge Repository of Aligarh Muslim University, would be extremely useful for the Institution in general and research community in particular. The proposed IR of AMU should serve as a one stop source *for the preservation and dissemination of institutional productivity*.**

F. Originality/value :- **Originality of this software architecture lies in the usage of OAI-PMH. 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.**

Keywords *Institutional Repositories, Aligarh Muslim University, OAI-PMH, Open Source Software, Dspace Software, Digitization, Self Archiving, Content, OAIS.*

I. INTRODUCTION

As centers for intellectual and scholarly research, academic and research institutions (whether in developed or developing countries) are expected to take an interest in the creation, dissemination as well as preservation of knowledge. In any given society, this is a very complex process, more so in developing countries where the economic, technological and institutional structures necessary to achieve the process are not well established (*Christian, 2009*). *Altbach* (1978) has noted that "knowledge dissemination is especially important in the third world context because the emergence of an independent intellectual life and some self-sufficiency in science is to some extent dependent on establishing the essential structure for dissemination of knowledge". Numerous academic and research institutions are sharing a challenging situation globally i.e., maintaining a large volume of material in various formats (printed, audiovisual and other types) and facing the difficulty of long-term digital storage and preservation of their data. Institutional repositories aim at the enhancement of the organization process of the rapidly developing volume of digital information that circulates in the environment of an academic institution. An institutional repository has the possibility of storing, providing access to, preserving and indexing a wide spectrum of digital material, which can be reused in various ways, and can be embedded in educational and socio-technical activities of the university (*Mechalis et al., 2012*). *Crow* (2002) defines Institutional Repositories as "digital collections capturing and preserving the intellectual output of a single or multi-university community. Lynch

ISSN No: - 2456 - 2165

(2003) described IR 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". Institutional repositories make possible wide dissemination of research outputs by the means of the internet. The growth of open access institutional repositories has been very remarkable in many developed countries (*Christian, 2008*). However, academic and research institutions in developing countries like India are still combating to be victorious over many challenging issues in effort to make their research outputs openly accessible by means of institutional repositories to the world wide audiences. Also Establishment of institutional repository especially in developing countries ensures that their national research becomes and contributes on an equal footing to the global knowledge pool (*Anuradha, 2013*). In the present era, one can't ignore the verity, about the importance of institutional repositories and their role in cooperating to academic excellence. So giving the precedence to the same; the present study has been designed to develop a proposed model of Institutional Repository that may be adopted as building block in all universities in general and in Aligarh Muslim University in particular.

II. OBJECTIVES OF THE PRESENT STUDY

The key objectives of the present study are:-

- To elaborate the skills and experiences in establishing an institutional repository at Aligarh Muslim University, Aligarh
- To enhance the visibility of the intellectual property and prestige of the institution to the outside world through IR.
- To discuss some key issues and challenges relating to the development of an institutional repository.
- To encourage other universities in developing institutional repository by highlighting the benefits of it in a university.

III. METHODOLOGY USED TO DEVELOP THE MODEL IR

The present study is based on pragmatic approach**.** Software is the essential part in developing and establishing Institutional Repositories. There are number of Software (Open Source & Commercial) packages available having diverse characteristics. To choose the best one according to the requirements, a comparative study of the available Open Source Software meant for developing OAR has been done while developing a proposed Model of Institutional Repository. Therefore, Dspace was selected, downloaded and installed, followed by metadata encoding and uploading of the collected material under the given circumstances. To populate the Repository, documents/publications were collected personally by approaching each and every Faculty member/Research scholar of the Faculty of Life Sciences, Aligarh Muslim University, Aligarh. The data collected from the scholars may be in two forms i.e. born digital (pdf) form and in print form. The born digital data was used '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.

IV. CONTENT IN THE PROPOSED IR

The Proposed Model of IR provides access to full text journal articles, abstracts, conference publications, abstract of theses, chapters in books, etc. and grey literature produced by the Faculty Members as well as Research Scholars.

V. SYSTEM REQUIREMENT FOR THE DESIGN AND DEVELOPMENT OF THE MODEL IR

The details of the System requirements and software versions used in developing a repository model are given as under:

A. Hardware:- To develop a repository, a high configuration server or host machine is desirable. Taking this into account all the factors, the system was used with following configuration I3-2100 CPU 3.10 GHz and Installed memory (RAM) 2.00 GB (1.70 GB usable), Hard Disk memory 500 GB.

B. Operating system:- Windows 7 OS was adopted to build up the repository.

C. Software:- Software is a vital part of a repository. For the present model, DSpace open Source software has been selected, because of its unique characteristics such as long-term physical storage and management of digital items in a safe, sound and proficiently managed repository. Besides, it has standard operating procedures such as backup, mirroring, refreshing media and disaster recovery.

D. Supporting Software Packages

- **Java JDK** jdk-6u35-windows-i586
- **Apache Ant** apache-ant-1.9.2
- **Apache Maven** apache-maven-3.1.1
- **Apache Tomcat** apache-tomcat-6.0.37
- **PostgreSQL** postgresql-9.3.1-1 windows

VI. DESIGNING MODEL IR: THE PROCEDURE ADOPTED

The whole procedure has been divided into different sections and phases for designing and developing the model repository of Aligarh Muslim University. These are discussed as below:-

ISSN No: - 2456 - 2165

A. First phase

In the first phase of the repository development the following procedure has been adopted.

- DSpace (DSpace-3.2-src-release) open source software was installed as per the instructions provided on the DSpace site.
- To complete the installation procedure, supporting software packages such as PostgreSQL, Ant, Maven, Apache Tomcat and Java JDK, were downloaded from their source sites.
- To get the software to perform functions as per the requirements of the repository, certain changes have been done in the installed package.

➢ *Workflow for the Repository Development*

The workflow of the repository is divided into the following sections which are described as below:

➢ *Organizational Structure of the Repository*

The organizational structure of the proposed AMU Repository system is described here as under:

There is 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 Theses& Dissertations, which are finally divided into collections. The organizational structure has been illustrated in the figure 1.

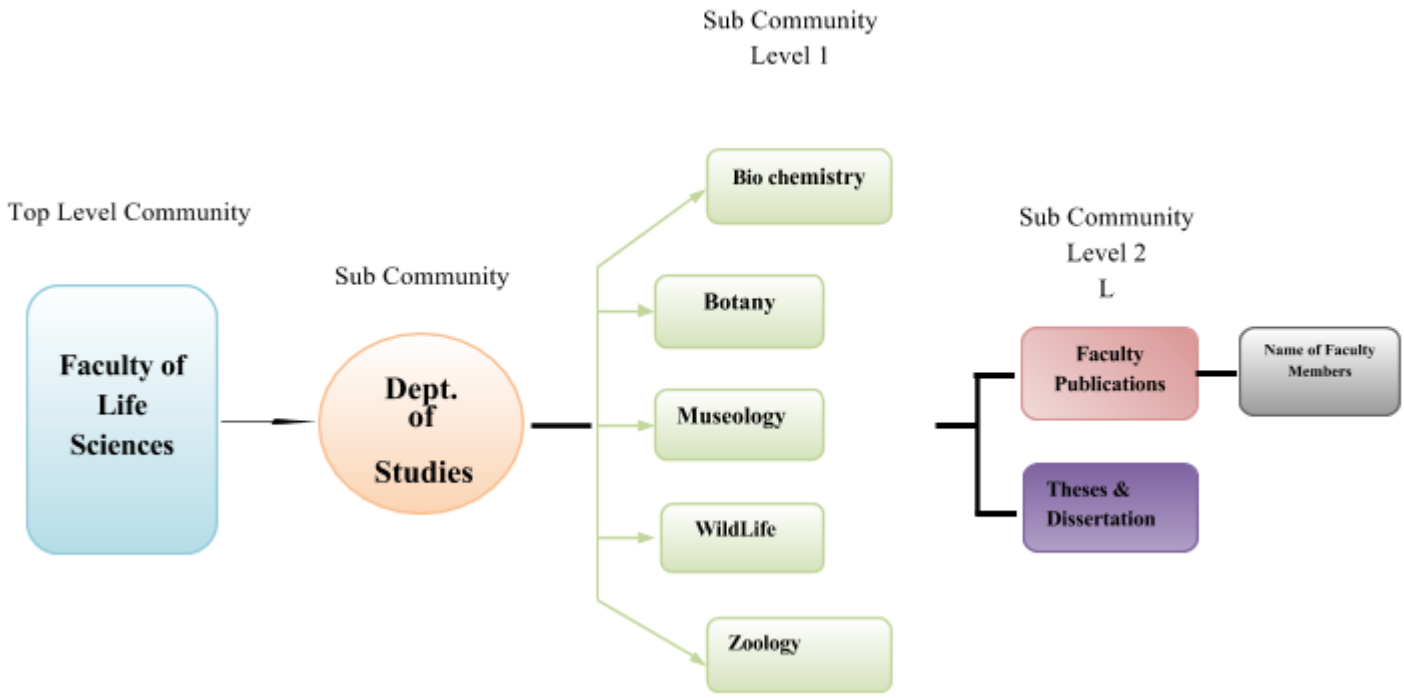

Figure 1: Organizational structure of the Model IR

According to *Hemmingway and Gough* (1998), "the functionality of a system cannot be considered distinct from the representation of information to its users". Processed data acquires their full meanings only when communicated to users as messages in their own languages and programs and the user interface are fundamentally associated (*Witten et al., 2005*). In alliance with this view, the dspace software has been customized included processes varying from the localization of the GUI and the preliminary texts in the homepages, to logo icons and the selection of colours in the homepage. An additional challenge concerning the representation of the information was related to the Semantics/Registry of the metadata. To make entry for 'Theses and Dissertations' clear and easy, one field of 'Guide' is added (Fig:2& 3).

ISSN No: - 2456 - 2165

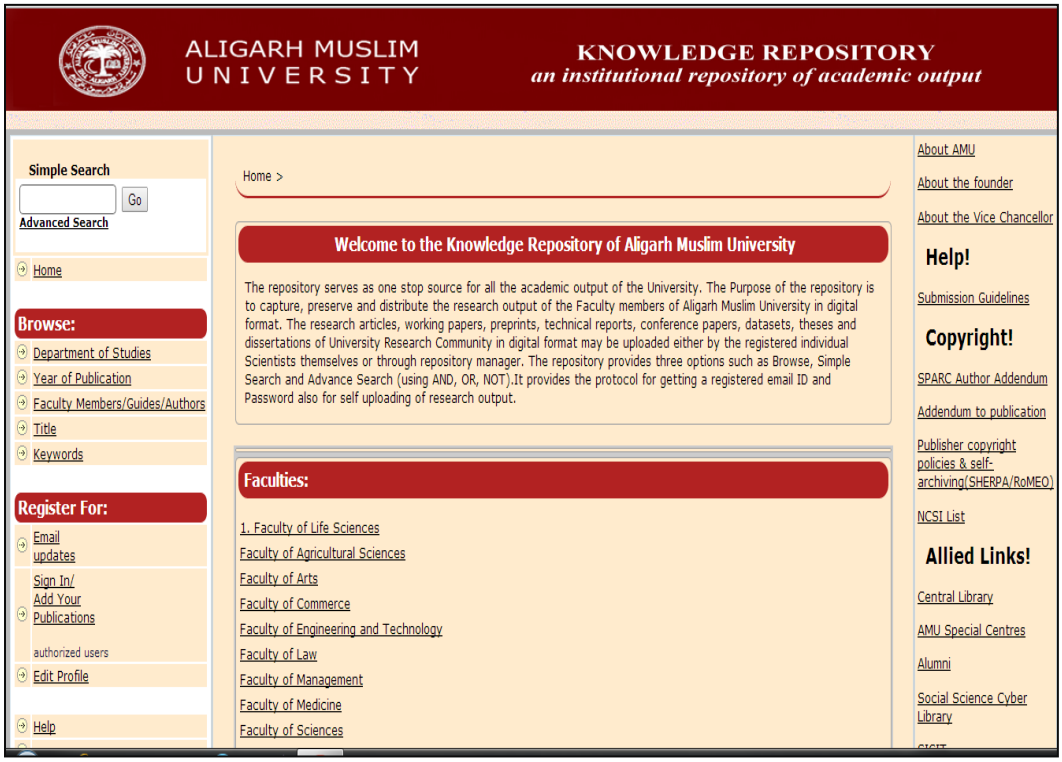

Figure 2: Knowledge Repository of AMU

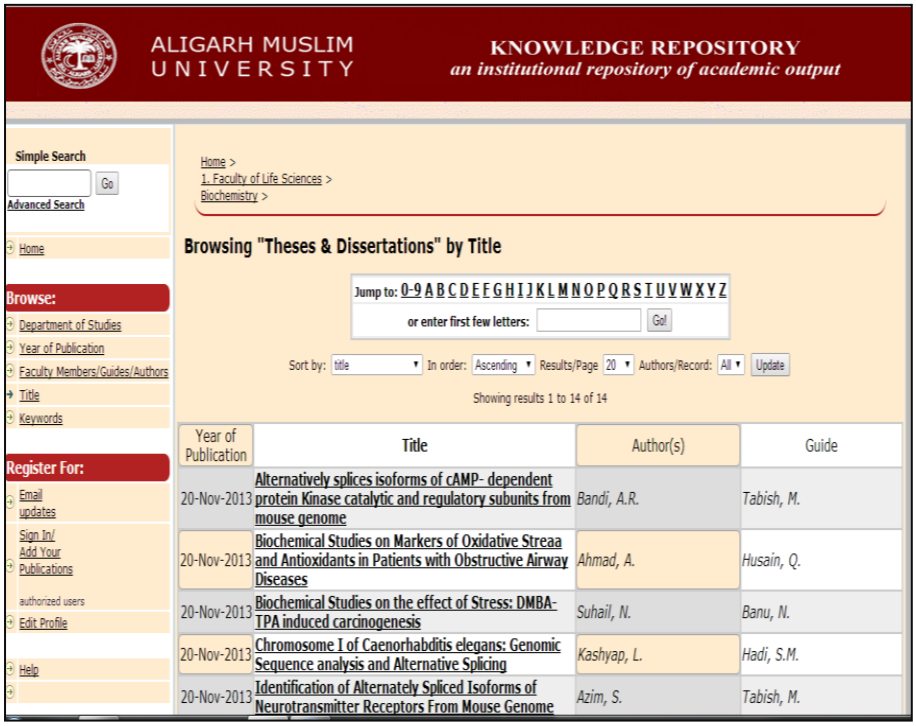

Figure 3: Modified Metadata Registry (Guide column)

ISSN No: - 2456 - 2165

B. Second phase

The second phase is concerned with the storing and indexing of the Faculty Publications and the creation of Metadata.

- *Protocol for Uploading of Contents*

In the process of development of repository, two different methods are used to create metadata and uploading of contents into the repository. Although the purpose of the present repository is to make the researchers self-reliant in uploading his/her research material, the same can also be done through the Repository administrator. The whole submission process is divided into seven steps that are represented by seven oval shaped buttons as shown in figure 4. The first three steps are used to describe an item, submitting in the repository while the remaining four steps are used for verification, License and uploading.

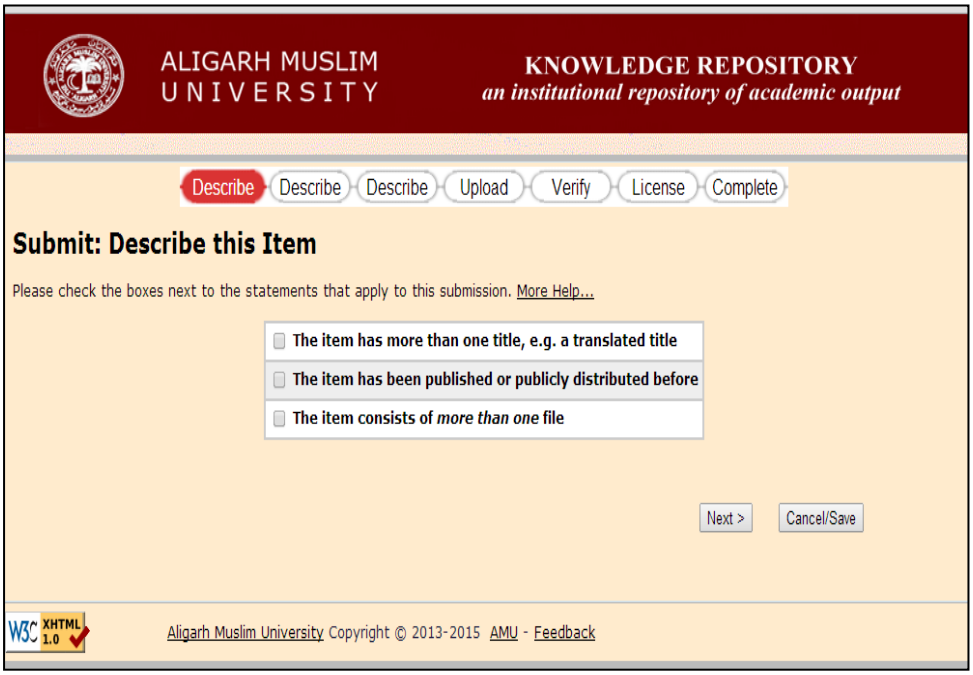

Figure 4: Uploading Process in IR

- *Deposition of Content*: This may be done by the methods described as below:-

1. Self Deposition/Archiving
2. Repository Administrator

- **Self Deposition/Archiving**: The repository has the provision of self Archiving/Deposition of contents by the registered users in requested collection(s). For this purpose, a user has to get himself/herself registered into the repository and may request the repository manager via email, to assign the requested collection(s) for self-uploading of the contents. Once the user gets the email and password, he/she can deposit his/her content into the repository himself/herself by following seven steps process as described under the heading **'Protocol for uploading of Contents'.**

- As soon as the content(s) are submitted by the Faculty Members/ Researchers in the proposed Knowledge Repository of AMU, it will enter the workflow process of the Knowledge Repository which has been earlier assigned by the Repository Administrator. Here, first the submitted content will go to the Metadata editor/ reviewer who will review the content and confirm metadata details / accept / reject the content. If rejected, the content will go back to the submitter for further improvements. The accepted content will pass on to the task pool of the collection administrator who will verify the required details and can accept or reject the submission. The rejected content will again go back to the submitter. Else, it will move to the Repository Administrator who will check copyright and other related issues and will decide that the content should be deposited into the Knowledge Repository or not. The workflow of the Knowledge Repository is shown in Figure 5.

ISSN No: - 2456 - 2165

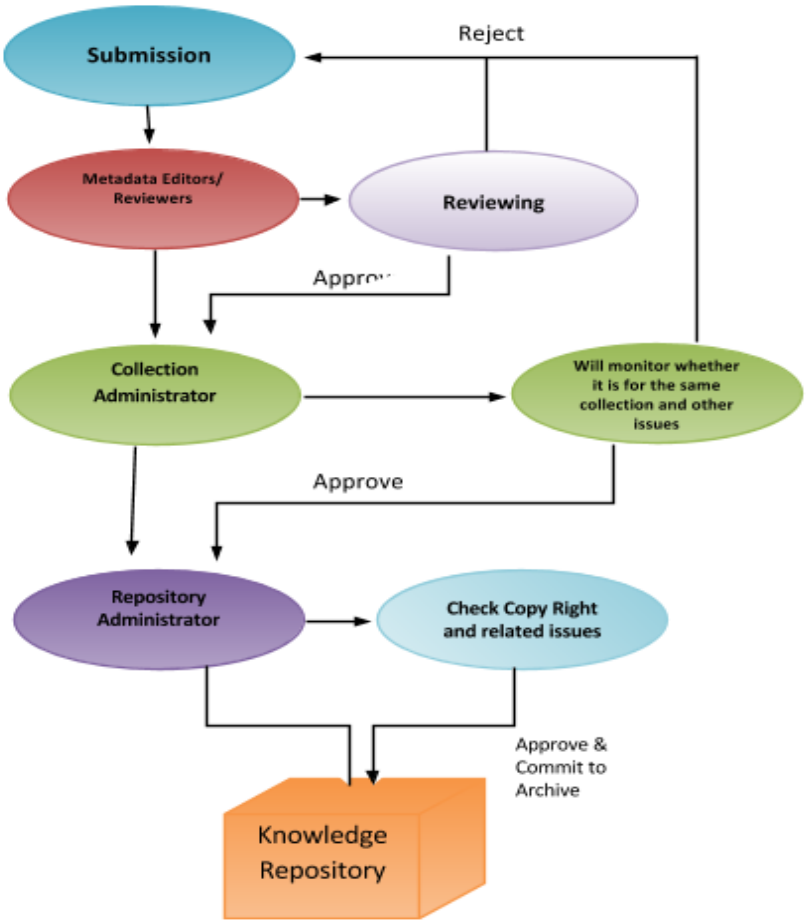

Figure 5: Work Flow in the Knowledge Repository

C. Third phase

The third phase is a testing phase and it involved the joint venture on the pilot/test server. During this stage, the functionality of the proposed model of the Institutional Repository of Aligarh Muslim University was tested and further upgraded by taking valuable feedback and suggestions from the Faculty Members as well as Research Scholars. 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.

- *Testing and feedback*

In the testing phase of the knowledge repository, the proposed model was put to use for Faculty Members and Research Scholars of Faculty of Life Sciences, AMU, Aligarh. There are a total of 61 Faculty Members and 220 Research Scholars (including PhD and M.Phil both) in the Faculty of life Sciences, AMU. A total of 281 questionnaires to collect data from the users were distributed in the Faculty, of which 210 questionnaires were taken back.

To get the feedback for the functionality of the proposed model of Institutional Repository various questions such as 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 were put up with the respondents.
The feedback given by the respondents can be drafted as follows:

- When the proposed model of IR was put to use by Faculty Members and Research Scholars from the Faculty of Life Sciences, AMU, Aligarh, it was astounding that a large number of users had not used an institutional repository before.

ISSN No: - 2456 - 2165

- 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.
- The users in both the categories were found to be entirely satisfied with the organization of various elements in model IR.
- When asked about their preference of publishing in proposed AMU Institutional Repository, as and when it is launched on permanent basis, a large number of users answered in positive.
- Regarding the mode of submission of articles in the proposed Institutional Repository of AMU, not many faculty members but the majority of Research Scholars preferred to use the method of "self-deposition."
- Users in both the categories were fully satisfied with the browsing and searching options provided in the model IR.
- As to the active links given in the model IR, majority of users were found to be fully satisfied.
- Also to the user friendliness of the Model IR, majority of users were found to be fully satisfied.
- Majority of respondents consented that the proposed model of IR when fully operational will increase the visibility of research and will be able to bring AMU at par with other top universities around the globe.
- A large number of respondents agreed that such a repository will enhance the number of citations and consequently impact factor.

VII. USERS OF THE PROPOSED REPOSITORY

There are three types of users in the repository

A. General user

A general user is a guest user who visits the repository and searches, browses and uses advanced search options. General/Guest users are granted permission for reading, downloading, sharing of full-text content for non-commercial purposes only. These users may or may not register themselves for accessing and downloading content from the Repository.

B. Repository user/Registered user

The registered users also 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 (Figure 6). The Repository user(s) can be a collection manager or a community manager in the repository, who have been given the privilege to manage particular collection or community or to self archive their content by the Repository administrator. The Repository users are also required to register with the repository.

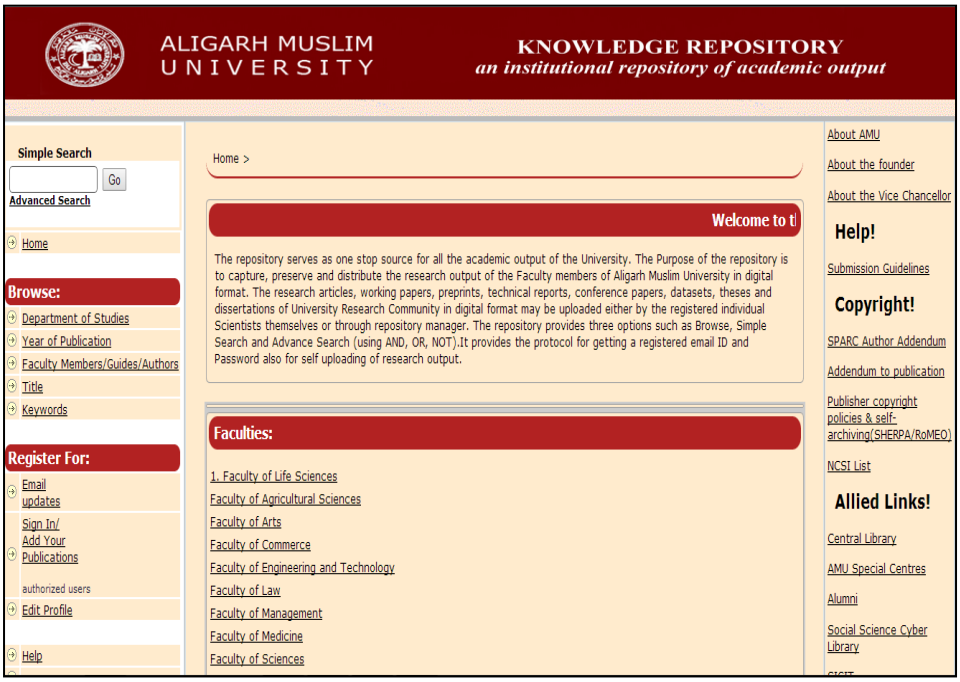

Figure 6: Registration in IR

ISSN No: - 2456 - 2165

C. Repository Administrator

Administrator of the Repository has administrative privileges to manage the content in the Repository along with the functions which a general user can perform. No one can perform any function until and unless the administrator allows him/her to manage the content in the Repository. The Repository Administrator is required to register with the Repository.

VIII. CONCLUSION

The paper presented the portrayal of work flow of the institutional repository, sighted as an example of the usage of technology in its social and organizational framework. It is assumed that the proposed model IR will serves as the starting point to build and develop IR not only for Aligarh Muslim University but it will provide solutions to develop such IRs to the other Universities. In a university that offers education to diverse clientele, the technological artifacts should be designed in relation to many interlinked constraints that can be fundamentally predictable within the context of a socio-cognitive engineering loom. 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. The additional value of the paper is that it may help Aligarh Muslim University and other universities and academic institutions to envisage challenges related to: diverse needs of various groups of users, current and future organizational policies, technological infrastructure and modify all that into useful input that enlightens the process of building an institutional repository service, which seems to be an ill-structured problem due to its veiled complexity. It is to be supposed that in near future the Aligarh Muslim University will have full-fledged Institutional Repository.

References

[1]. Lynch,C.A.(2003). "*Institutional repositories: Essential infrastructure for scholarship in the digital age*", ARL: A bimonthly report,226:Feb,1-7, available at: http;//www.arl.org/newsltr/226/ir.html (accessed 3 December 2015)

[2]. Crow, R. (2002). "*The Case for Institutional Repositories: A SPARC Position Paper*", available at: http://www.sparc.arl.org/sites/default/files/media_files/instrepo.pdf (accessed 4 November 2015)

[3]. Anuradha, K.T. (2013). "Design and development of institutional repositories: a case study. The international information", Library Review, Vol. 37 No. 3, pp. 169-178.

[4]. Altbach P. G. (1978). "Scholarly publishing in the third world", Library Trends, Vol. 26 No. 4, pp. 489-504.

[5]. Christian, G.E. (2009, January). "Issues and challenges to the development of open access institutional repositories in academic and research institutions in Nigeria" (Research Paper).Ottawa: IDRC, available at: https://idl-bnc.idrc.ca/dspace/bitstream/10625/36986/1/127792.pdf (accessed 4 December 2015)

[4].Mechalis, X., Anna, M. Dionisis, K. George, B., Dimitris, G. and Dionisis, Z. (2012). "The development of an institutional repository service for a large distance education university", in *proceedings of the EADTU 25th Anniversary Conference 2012, The role of open and flexible education in European higher education systems for 2020: New models, new markets, new media*, Pafos, Cyprus, pp. 327-337, available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.413.9977&rep=rep1&type=pdf (accessed 29 November 2015)

[5]. Hemingway, C. and Gough, T. (1998). "A Socio-Cognitive Theory of Information Systems", Information Systems: Current Issues and Future Changes, pp. 275-286.

[6]. Witten, I., Bainbridge, D., Tansley, R., Huang, C. and Don, K. (2009). "Stone D: A Bridge between Green-stone and Dspace", D-Lib Magazine, No. 9 (online), available at: http://www.dlib.org/dlib/september05/witten/09witten.html (accessed 3 December 2015)