ISSN-2230-9659

An Evaluative Study on the Open Source Software for Building Digital Libraries

Pooja Sharma
Librarian, Shri Krishan Institute Of, Engineering and Technology (SKIET) Kurukshetra (Haryana) – 136118

ABSTRACT:-

Information technology has played an important role in library and information science. Due to the developments in information technology, now, it is possible for libraries to provide several new services to the library users along with traditional services. Libraries are now able to provide information in print form as well as in digital form.

During 1980s libraries started automating their bibliographic databases and during 1990s digital library projects were initiated. As on today, lot of developments have taken place in digitizing print media. At national and international level several big funding projects have been initiated to digitize valuable material available within the libraries for the preservation as well as for providing wider access to the collections through latest technologies.

Digital libraries have been making their roots in the library profession as a separate discipline and many conferences, workshops and seminars are taking place in the area of digital libraries. These conferences are covering different topics under digital libraries such as collection development and organization, user studies, digital library architecture, usability studies, search and retrieval, digital library software or providing value added services to end users.

ISSN-2230-9659

Digital libraries are becoming popular and are becoming one of the important activities of any organization. The rapid growth in computing networks, databases and public awareness have contributed to a hot topic of today such as digital libraries, digital archives, institutional repositories or digital repositories.

INTRODUCTION:

According to Deegan, M. and Tanner, S., definition of a digital library is

1. A digital library is a managed collection of digital objects.

2. The digital objects are created or collected according to the principles of collection development.

3. The digital objects are made available in a cohesive manner, supported by services necessary to allow users to retrieve and exploit the resources just as in the case of other library materials.

4. The digital objects are treated as long term stable resources and appropriate processes are applied to them to ensure their quality and survivability.

Digital Library Foundation defined digital library as “Digital libraries are organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.” This definition involves three key components, which constitute the theoretical framework of digital libraries such as: people, information resources and technology.

Another workable and widely used definition of digital library provided by Lesk is “a collection of information which is both digitized and organized”.

A review of the above definitions indicate that digital libraries are organised in a special manner with the help of computers and are made available over a network, with different procedures such as, to select the material for adding in the collections, to organize collections, to make collections available to end users with value added services and to archive the material for long term access.

During 1990s, the research development and practice related to digital libraries took off in a large scale. Researchers focused on digital libraries as contents collected on behalf of user communities, while librarians focused on digital libraries as institutions or services. Many new initiatives came through the

ISSN-2230-9659

research sponsored by the US National Science Foundation (NSF) and UK Joint Information Systems Committee (JISC).

Digital Library Initiative (DLI) projects in the USA and the projects in UK have played a key role in the development of digital libraries. In addition many digital library projects are currently under way in Australia, Asia, Europe, Africa and Latin America. While some of them have their own funding, others are funded under digital library-specific funding initiatives.

Today many other groups as well are involved in the expansion of digital library technologies and techniques such as European Union, Association for Computing Machinery (ACM), the Institute of Electrical and Electronics Engineers (IEEE), The International Federation of Library Associations (IFLA), the American Library Association (ALA), the Coalition for Networked Information (CNI), and the Digital Library Federation (DLF).

During the past decade thousands of digital libraries in a variety of forms were built glob- ally and are functioning operationally, with more to come. Hundreds of research projects were then devoted to many aspects of digital libraries in many countries, and more are reported each year.

A fully developed digital library environment involves the following elements:

1. Initial conversion of content from physical to digital form.

2. The extraction or creation of metadata to assist in object viewing, management, and preservation.

3. Storage of digital content and metadata in an appropriate multimedia repository.

4. Client services for the browser, including repository querying and work flow.

5. Content delivery via file transfer or streaming media.

6. Patron access through a browser or dedicated client.

Digital libraries thus provide new technological platform for implementing functionality of traditional library systems by making them much more powerful. Digital libraries developed today are based on innovative web technologies such as Semantic Web, Ontology Specification, Database Technologies, XML databases, text retrieval in different languages etc.

ISSN-2230-9659

While using the technology, it has also placed couple of challenges in front of the librarians such as which hardware/software to be used for organizing scanned digital collections or born digital collections, how to maintain these collections over a long term, what are the other aspects which needs to be considered before bringing digital collections on the Internet/Intranet of Library Web Sites.

REVIEW OF LITERATURE:

In the initial stage of research an extensive literature search was carried out on the following databases on key terms such as Digital Libraries, Evaluation of Digital Libraries, Evaluation of Open Source Digital Library Software and Open Source Digital Library Software for understanding the research covered by earlier researchers. The searches were carried out on Internet as well as on Online bibliographic and full text databases. Contacts with the colleagues, email discussion lists and other appropriate sources were also used to gain the knowledge of actual work done so far in this research area. While reviewing the material, the emphasis was given on providing some insight into the current level of activity in the evaluation of open source digital library software.

SOURCES OF LITERATURE

1. LISA: Library and Information Science Abstracts is an international abstracting and indexing tool designed for library professionals and other information specialists. LISA currently abstracts over 440 periodicals from more than 68 countries and in more than 20 different languages. The abstracts are covered from 1969 till present. Over 296,127 records are there on CSA website as of March 2007.

2. ACM Digital Library: The ACM digital library covers major digital library science conference proceeding named as Joint Conference on Digital Libraries since 1997. Hence this database was used in depth to know the latest research carried out in Digital Libraries.

3. Google Scholar as well as General Google: Google Scholar covers bibliographic as well as full text database of all articles published from all journals and google covers presentations, articles available from individual web sites for downloading.

4. Science Direct: Covers major library science journals such as Information Processing & Management, Library and Information Science Research etc.

5. Springer Link Services: Covers all articles from the journal International Journal on Digital libraries which is a European Journal.

ISSN-2230-9659

6. Emerald Database: Covers major library science journals such as Journal of Documentation, Library Hi Tech, Library Management, Library Review, The Electronic Library, Online Information Review, Program:electronic library and information systems

While undertaking the review of the literature related to the key focus areas of the present study the researcher has concentrated on locating material to provide contextual information and background. Information has also been identified which examines the relationship between theory and practice in the areas of evaluation of open source digital library software.

RESEARCH METHODOGY:

Paul Jones emphasized on a few key components of most open source projects such as open contribution, user centered feedback & distribution, and building of user communities. These trends have been identified as significant reasons to shift to open source initiatives.

A couples of digital library software are available for the librarians to build, maintain, organize and make their collections available over Internet for end users. These OSSDL are available on Internet under Open Source License for helping librarians to manage dig- ital collections and make them available on Internet/Intranet. To name a few Greenstone, DSpace, Fedora, EPrints, DoKS, MyCoRe, ARNO, CDSware are available on Internet un- der Open Source License form. OSS4Lib site maintains major open source software used in libraries for all applications that are carried out in libraries.

For years institutions have relied upon commercial vendors to create and maintain the software that they use – from small digital library projects to large enterprise financial systems but now OSS provide a viable, and preferable alternative to vendor provided software.

Many of the OSS-DL applications such as Fedora software, DSpace software and Greenstone software have come from grant funded projects. The academic community all over the world looked towards these open source software as an alternative to commercial solutions.

All these software have their own advantages and disadvantages. Since there are couples of OSS available for building digital libraries it is difficult for the libraries to decide which software is most suitable for them in a given set of environment.

As the use of open source software (OSS) has taken off over the past decade, there has been increasing interest in the potential of open source to address longstanding concerns in the library and

ISSN-2230-9659

information science regarding the cost and performance of commercial software products. A common view is that existing proprietary options do not have the features required or allow for cost-effective customization. There is an active open source community which has been constantly trying to provide a better solutions in all areas.

So far very few studies have been carried out in evaluating OSS-DL (as mentioned earlier) and these studies have not covered all software that are available under Open Source license terms and conditions as well as have not covered extensive evaluation criteria. Hence there was an urgent need to define extensive evaluation criteria and carry out in depth evaluation of all OSS-DL available free on Internet and preferably under open source terms and conditions and find different features supported by these software.

NEED OF THE STUDY:

Digital libraries give us opportunities we never had with traditional libraries or even with the web. The advent of open source and proprietary software has enabled universities to develop web based digital repositories.

In the fast changing networked environment librarians have more responsibilities. Librarians have to tackle print media as well as digital media. Handling digital media demands more up to date knowledge of the subject as well as technology.

Though librarians try to digitize important collections available with them, the further biggest challenge is how to organize the born digital collections as well as digitized collections which are spread around in all organizations.

During 1998, open source trend came into libraries and libraries started making use of these OSS tools for different tasks carried out in libraries. OSS tools are helping libraries to overcome problem of high budget allocations for buying commercial solutions. Lot of research is going on in using OSS based applications which are useful in libraries. In recent years, there have been steep rises in the number of applications that are available for the libraries to automate library catalogues, MARC editors, digitization work flow management software, web page designing software etc.

HYPOTHESES:

1. Many of the Open Source Digital Library software (OSS-DL) are not in conformity with standards.

ISSN-2230-9659

2. Digital preservation is not really yet been addressed by many of the OSS-DL software.

3. Interoperability, scalability are highly questionable aspects of much OSS-DL software.

OBJECTIVES OF THE STUDY:

The objective of the present work is to evaluate major Open Source Digital Library Soft ware (OSS-DL) available free and preferably under Open Source License terms and conditions. The main objective or goal of the present work can be further divided into the following sub-goals or secondary objectives:

1. To identify a list of major Open Source Digital Library Software available for creating digital libraries on Internet.

2. To identify the evaluation studies carried out so far on Evaluation of Open Source Digital Library Software (OSS-DL).

3. To design a set of evaluation criteria for evaluation of selected Open Source Digital Library Software (OSS-DL).

4. To install selected candidate software for carrying out detailed evaluation study on a test bed environment.

5. To evaluate each selected software against the set of designed criteria such as content acquisition, content management, metadata submission and support, classification, information search/retrieval, access control, privacy and management, authentication and authorization, interoperability, ease of deployment of each software, security, scalability, backup/restore facility, user friendly interface, usability, trans- action log analysis, copyright issues, load balancing, personalization, visualization, digital preservation etc.

6. To find out at what level digital preservation is supported in each software.

7. To enlist each criterion and against that evaluate each OSS-DL capabilities, along with its advantages and disadvantages which will help librarian to decide which soft- ware to be used for a given set of environment with a given set of features.

SCOPE OF THE STUDY:

ISSN-2230-9659

All over the world libraries have started creating digital libraries of important documents available with them. For managing digital documents, libraries have either commercial information management systems or open source software.

Use of OSS for libraries has grown in the last decade. There are large numbers of OSS software available for carrying out different tasks in the libraries. Since the source code is available it is easier to customize the software as per the end users requirements. Considering the fact that OSS is free for use, it is of great help to developing country libraries. The implementation of OSS in libraries helps to improve library services and collections.

There are couples of OSS-DL available today. Each software satisfies a particular need. Since each of the software has different configuration, have used different standards, satisfies different needs for creating DL, it was felt necessary to study all the major currently available Open Source Software for Building Digital Libraries with various factors which will help libraries especially from developing countries what features are supported by each software, what are the installation steps, which software has sustainability to stay in future, what are the advantages/disadvantages of using each software, how many users are using the software etc.

The present study covers only those software which are available for creating digital libraries. This study does not cover any other library related open source software (for e. g. Library automation software such as Koha, New Gen Lib, CDS-ISIS etc.) though they are available under OSS license.

The scope of the study is further extended to evaluating current stable versions of all software that are available on Internet during the course of work. The scope is also further restricted to successful installation of each software which are taken for the present study.

The primary objective of this study is not to present which is the best software as compared to the other one. It does not attempt to rank any software and indicate the best software. In fact, such a comparison many not lead us anywhere because each software has its own objectives behind developing it. However, what features are supported by each software may help us to understand range of facilities available in each software and which software best suits for a particular type of library will certainly be helpful.

CONCLUSION:

ISSN-2230-9659

The digital media as on today is considered as a stable media along with the print media. Libraries are creating DL of their valuable collections. Many live repositories are available today and libraries are using OSS-DL tools effectively to build institutional repositories/digital libraries/digital repositories. In developing countries where funds are less OSS have been found very useful. OSS has much potential in library science and there are number of projects that demonstrate its viability. It gives library staff an option to be actively involved in OSS development projects and this involvement takes many forms, such as reporting bugs, suggesting enhancements, usability studies or testing new versions of OSS.

Considering these opportunities it was felt to evaluate available OSS-DL in detail and identify features supported by each software and evaluate them against a set of defined evaluation criteria. This study provides insightful information to the user community about which software best suits in a given set of environment. This analysis also finds out pros and cons of each OSS-DL which will give guidelines to the libraries for deciding what type of OSS-DL can be used for a specific type of digital collection in a given environment setting.

REFERENCES:

- Deegan, Tanner. *Digital futures: strategies for the information age*. London, Library Association Publishing, p.22.
- A working definition of digital library. http://www.diglib.org/ about/ dldefinition.htm
- Lesk, M. E. *Practical digital libraries: books, bytes and bucks*, San Francisco, Morgan Kaufman, 370 p.
- McCray, T. & Gallagher, M. E. Principles for digital library development.
- *Communications of the ACM, 44(5)*, p.49-54.
- *Digital library technology trends*, Sun Microsystems, Inc., August 2002, 37 p.
- Kogalovsky, M. R. Digital libraries: ongoing development. *Programming and*
- *Computer Software, 28(4)*, p.185. http://www.maik.ru/ abstract/ procom/2/ procom4_ 2p185abs.htm.
- http://www.sourceforge.net.

ISSN-2230-9659

- Anctil, Eric & Beheshti, Jamshid. Open Source Integrated Library Systems: An Overview.
- Raymond, Eric S. *The cathedral and the bazaar*, O'Reilly, 255 p.
- Morgan Possibilities of open source software in libraries, *Information*
- *Technology and Libraries, 21(1)*. http://www.lita.org/ala /lita/litapublications/ital/ 2101morgan.cfm.
- Glass, R. L. A socio political look at open source. *Communications of the ACM, 46(11)*, p.21-23. http://www.oss4lib.org.
- Morgan, E. Open Source Software in Libraries: a workshop. http://infomotions.com/musings/ossnlibraries-workshop/.
- Burger, T. N. Libraries facilitate open access to information with open source software.
- Krogh, G. von & Hippel, E. von. Special issue of open source software development. *Research Policy, 32(7)*, p.1149-1291.
- Surman, M. & Diceman, J. Choosing open source: a guide to civil society organizations.
- http://www.commons.ca/ articles/ fulltext.shtml?x=335#whatis.
- Breeding, M. An update on open source ILS. *Information Today, 19(9)*, October, http://www.onlineinc.com/ it/oct02/ breeding.htm.
- http://www.csa.com
- http://portal.acm.org
- http://scholar.google.com
- http://www.sciencedirect.com http://www.springerlink.com
- http://www.emeraldinsight.com