

Adoption of Open Source Digital Library Software Packages A Survey

Sanjo Jose

National Centre for Science Information
Indian Institute of Science
Bengalooru -12

sanjo@ncsi.iisc.ernet.in

Original Manuscript submitted to Convention on Automation of LIBraries in Education and
Research Institutions (CALIBER) 2007

Abstract

Open source digital library packages are gaining popularity nowadays. To build a digital library under economical conditions open source software is preferable. This paper tries to identify the extent of adoption of open source digital library software packages in various organizations through an online survey. It lays down the findings from the survey.

Keywords: Open Source; Digital Library; Software; DSpace; EPrints; Fedora; Greenstone

1. Introduction

Digital libraries enable the creation, organization, maintenance, management, access to, sharing and preservation of digital document collections. Digital libraries are being created today for diverse communities and in different fields: such as education, science, culture, development, health, governance and so on. The purposes of indigenous culture like virtual museums. It may be also to accumulate locally produced collections of information or to build a repository of the scholarship of an organization. Digital libraries may differ from organization to organization, which implement them. It may be dissemination of information, preservation and propagation of indigenous culture like virtual museums.

2. Digital Library Software Packages

For the construction and administration of a digital library one needs digital library software. Many commercial digital library software packages are available today. But they are too costly to be afforded by average libraries. These proprietary software packages also require high installation costs followed by annual maintenance costs (AMC) and/or updation costs. A remedy for this situation is the adoption of open source software for digitization purposes. With the recent availability of several open source digital library software packages, the creation and sharing of information through digital library collections has now become an attractive and feasible proposition for library and information professionals around the world. To build a digital library under economical conditions open source software is preferable. This paper tries to identify the extent of adoption of open source digital library software packages in various organizations through an online survey.

3. Methodology

A search in the Internet for the availability of different digital libraries and digital library software packages retrieved numerous subject oriented, institution oriented and mission oriented digital library projects and a large variety of software packages for digital object management. About twenty open source digital library software packages were identified. As an exhaustive study of all them was not possible due to time constraints, four most popular digital library software packages were selected for the study. The selected software packages are DSpace, Eprints, Fedora and Greenstone.

A survey was conducted using online questionnaires to understand the extent of adoption of open source digital library software packages in various institutions in different parts of the world especially in India. A questionnaire was prepared with nine questions (questionnaire is given in appendix). The URL of web page containing questionnaire was mailed to various libraries in different parts of the world, by identifying their e-mail addresses. A request for filling the questionnaire was also posted on mailing lists of DSpace, EPrints, Fedora and Greenstone and on various listservs of library and information science communities of India like LIS-forum, Digital Library Research Group and IndiaLIS.

The respondents include Bodleian Law Library of Oxford University, The British Library, University of Illinois at Urbana-Champaign, Oregon State University, University Of Bristol, University of Texas, DRTC, NCSI, etc.

4. Hypothesis

The study started with the hypothesis that the number of installations of four open source digital library software packages DSpace, Eprints, Fedora and Greenstone has no significant difference.

5. Analysis of Online Survey

Altogether 79 responses were yielded in the survey. But thirty-five of the respondents were removed, as they were not quite aware of the digital library software packages. The remaining responses were from the people in various capacities like Librarian, Information Systems Librarian, Liaison Librarian, Associate Professor, Assistant Professor, Lecturer, Scientist, Technical Director, System Administrator, Library Systems Developer, EPrints Assistant, Institutional Repository Manager, Historical/Cultural Specialist, Program Manager, Project Assistant etc.

5.1 Distribution of Software Packages

Some institutions use more than one software package for their archival requirements. So the total number of software installations exceeds the total number of respondents.

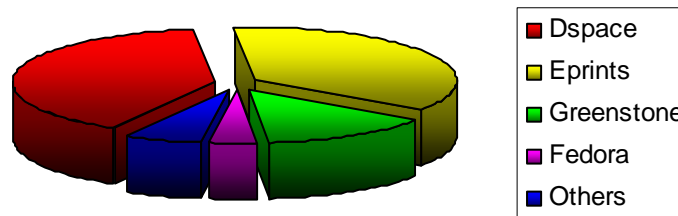


Figure 1
Distribution of software packages according to the online survey

Among the total 50 software installations included in the survey, DSpace with 20 installations seems to be the most popular choice among the digital library software packages. EPrints is at the second place with 18 installations. Greenstone has 7 installations and Fedora has got two installations. Software packages falling in 'other' category is avoided in the further analysis, as it does not come under the purview of the study.

5.2 Country-wise Distribution of Software Packages

Country	DSPACE	EPrints	Fedora	Greenstone	Total
India	13	7	-	4	19
USA	4	2	1	2	9
UK	1	5	-	-	6
Germany	1	1	1	-	3
Argentina	-	-	-	1	1
Denmark	-	1	-	-	1
France	-	1	-	-	1
Netherlands	1	-	-	-	1
Spain	-	1	-	-	1
Total	20	18	2	7	47

Table 1
Country-wise distribution of software packages according to the online survey

From table 1, it is evident that DSpace is the most widely used digital library software in India, with 13 installations. EPrints and Greenstone are also used in India. But digital libraries based on Fedora could not be found. US institutions are using

DSpace more, may be because of its American origin. Similarly, UK based institutions use EPrints more than any other software. Its cause may be attributed to membership of these institutions in various JISC funded consortia where they use Eprints for their repositories. Though India dominate in the following table with most number of installations, it does not mean all the digital libraries are concentrated in India. The questionnaire was more distributed among Indian professionals and so their participation was more in the survey.

5.3 Availability of Digital Libraries on WWW

Although many institutions have digital libraries, all of them are not available on World Wide Web. Among 47 digital libraries included in the survey, only 27 are accessible through Internet. Rest of the digital libraries is confined to the members of institutions alone. To allow open access these repositories of knowledge should be online.

5.4 Category-wise Distribution of Institutions

Among the total number of respondents of the survey (44 institutions), 31 were educational institutions comprising universities and colleges. 10 of the respondents were research institutions. Thus educational and research institutions were the majority of the respondents. Two of the respondents are consortia and one was a cultural organization based in Argentina. The category-wise distribution of institutions responded to the online survey is represented in figure 2.

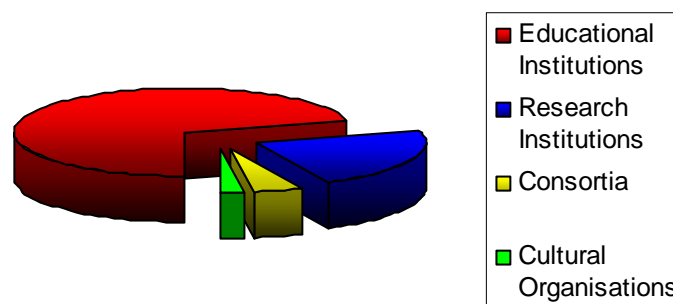


Figure 2

Category wise distribution of institutions responded to the online survey

5.5 Category-wise Distribution of Software Packages

Among the 44 digital libraries identified in the survey, 31 are maintained by educational institutions like universities and colleges. Research institutions are at the second place with 10 digital libraries. Educational institutions seem to prefer DSpace

and EPrints, may be due to its academic community specific architecture. Both are specially designed to preserve the scholarly output of organizations. The only cultural organization responded to survey uses Greenstone. Category-wise distribution of the software packages is given in table 2.

Category	DSpace	EPrints	Fedora	Greenstone	Total
Educational institutions	13	11	2	5	31
Research institutions	6	3	-	1	10
Consortia	-	2	-	-	2
Cultural organizations	-	-	-	1	1
	17	16	2	7	44

Table 2
Category-wise distribution of software packages according to the online survey

5.6 Some Revelations from the Survey

Though the questionnaire was sent to more than 500 e-mail addresses, it yields only one tenth of desired response. Thirty-five respondents have to be removed from the survey just because they do not have enough knowledge about the digital library software packages. It has to be said that all these respondents were from India only. Most of them gave the details about automation software packages used in their libraries. It is evident that many of the practicing qualified librarians also do not know the difference between automation software and digital library software. They call their libraries as “digital libraries” if it is automated. Even one of the respondents (who is the Head of Library and Information Science Department in a University in India) went to the extent of asking about the omission of Koha in this survey. Koha is open source library automation software.

The survey thus emphasizes the need of making library professionals (and even teachers) making aware of different concepts and technologies in library and information science and related disciplines. Library professionals should be up-to-date at least in their discipline as they have access to modern information sources.

6. Conclusion

DSpace is the most popular among the digital library solutions available in the open source domain. Eprints is also widely used. Educational institutions dominate in the use of these packages. Though many institutions have implemented digital libraries, only about half of them are available online. Open access of knowledge is possible only if these repositories are made online. India is benefiting well from the open source movement.

Appendix: Questionnaire used in online survey

1. Name
2. Designation
3. Organization
4. E-mail
5. Please mention which software is used for the digital library/repository in your institution?
 - DSpace
 - EPrints
 - Fedora
 - Greenstone
 - Other (Please specify)
6. Please provide the URL of the digital library/repository? (If it is available over WWW)

About the author



Sanjo Jose has done his Master of Library and Information Science from the Department of Library and Information Science, University of Kerala, Thiruvananthapuram in 2006 and now is working as Information and Knowledge Management Trainee at National Centre for Science Information, Indian Institute of Science, Bengaluru.