
OPEN SOURCE SOFTWARE: A COMPARATIVE STUDY OF GREENSTONE AND DSPACE

S N SINGH

R K NGURTINKHUMA

PANKAJ KUMAR SINGH

Abstract

Open Source Software (OSS) is software available free of cost with source code for anyone to use. Since last few years, it is gaining attention of library and information professionals. Digitization involves huge money to create and maintain and the OSS appears to be a means to reduce it. Among these, DSpace and Greenstone are becoming more popular in India and abroad. This paper deals with the comparison of these two popular OSS from various points of view. The comparative table may help the professionals who are planning to create a digital library/ institutional repository.

Keywords : Open Source Software/ Greenstone/ DSpace/ Institutional Repository

1. Introduction

Library profession is moving away from traditional library towards the creation and maintenance of a digital library. Information contents that were confined to traditional formats like books, journals, maps, sound recordings are getting increasingly available in diverse digital formats. New formats, being the core elements of digital collection, have emerged such as multimedia, hypertext, dynamic pages, interactive video, etc. Each format poses distinct challenges for its preservation and access. Capturing, storing, indexing, preserving and redistributing content with ease of use and web-based user interface are some of the core challenges of any digital library that are being faced by the library professionals.

Open Source Software (OSS) has appeared as an open key to the digital library and creation of the Institutional Repositories (IR). Over the last couple of years, the Open Source Software has grown tremendously in scope, popularity and use. It has gained the attention of library professionals and researchers because of its features including free availability, as cost is major factor behind digitization activities.. OSS refers to the software which is freely available for use. The idea is that every one is allowed to study or use, and thus to change and improve the software which is being distributed. Usually, open source is distributed free, either with no restrictions on use or under the GNU

Public License (GPL). This software makes the source code available to users to view, amend and adapt. Another important feature is access to its source code, which allows the participation of participants and users. A source code consists of a set of instructions which are translated into binary codes so that the computer can understand the instructions and allow the user to interact. Windows, Unix, Oracle, etc are the examples of closed source programs where the source code is secret to users. This can be made more useful by essential modifications and adaptations according to needs of individual library. It may be an alternative to software that is commercially available, usually expensive and difficult to customize.

The Website www.opensource.org defines “Open source promotes software reliability and quality by supporting independent peer review and rapid evaluation of source code. To be certified as open source, the license of a program must guarantee the right to read, redistribute, modify and use it freely.”

1.1 Features of OSS

Open source software has the following characteristics:

- OSS is freely available to users, usually through a license. Source code is free to change, modify and redistribute it.
- OSS is normally created and maintained by developers crossing institutional and national boundaries, collaborating by using internet-based communications and development tools.
- Successful applications tend to be developed more quickly and with better responsiveness to the needs of individual library or institution who can readily use and evaluate open source software applications because they are free.
- Intellectual Property Rights to open source software belongs to everyone whoever is involved.
- It does not involve the profit.
- Cooperation and quality are the motives of open source software developers.

1.2 History and Development of OSS

The first open source software came into existence in 2000 with release of eprints.org. Since then the Open Archives Initiatives (OAI) has gained momentum and getting the attention of library and information science professionals. A number of OSS have appeared in recent past and are available on Internet. These include:

- Eprints,
 - Greenstone,
 - DSpace,
 - Ganesha Digital Library,
-

-
- CDSware, (CERN Document Server Software)
 - ROADS, (Resource Organization And Discovery in Subject-based Services)
 - iVia,
 - Phronesis,
 - I-Tor
 - FEDORA (Flexible Extensible Digital Object Repository Architecture)

Amongst the softwares mentioned above, DSpace digital repository system, Greenstone Digital Library, Eprints, and Ganesha Digital Library are most popular. The Institutional Repositories in India are using open source software namely DSpace, Greenstone, EPrints, etc.

2. GREENSTONE

The greenstone digital library software is one of most important and famous open source software. This software is developed and distributed as an international cooperative effort established in August 2000 among three parties namely New Zealand Digital Library Projects, University of Waikato, UNESCO and the Human Info NGO, Belgium. It is developed and distributed in cooperation with UNESCO and the Human Info NGO. Greenstone v2.71 is the latest version. This version will serve existing collection without any change. Greenstone is a suite of software for building and distributing digital library collections. It is open source, multilingual software for the construction and presentation of digital information. It provides the facility to build collections with effective full text searching and metadata-based browsing. Browsing facilities are attractive and easy to use. This project that works with UNESCO and NGOs has established a worldwide reputation for building digital collection and making it widely available, free of charge to developing nations and on a cost –recovery basis to others. The purpose of the Greenstone software is to empower users, particularly in universities, libraries and other institutions to build their own digital collections. It provides a new way of organizing information and publishing it on Internet or on CD-ROM. The complete Greenstone interface and all documentation is available in five core languages: English, French, Spanish, Russian and Kazakh, interface can also be extended to other languages. Collections can include text, audio, video and pictures using an easy to use tool called the Collector. Documents in the collection can be in any language. Even GSDL interface is available in many languages. Using Greenstone, collections can easily be maintained and effective full text searching and metadata based browsing facilities are extended.

2.1 Technical Platforms

Greenstone software has been tested on all versions of Windows (95/98/NT/2000/XP), UNIX, Linux, and Mac OS-X. Greenstone v2.70 or v2.71 can be downloaded and installed easily. Windows users need not any configuration when downloading it on their computer terminals or laptops. It is very easy to download, test and use.

2.2 Interoperability

Greenstone is completely interoperable as it uses contemporary standards. It has the ability to interact with two or more systems/ metadata schemes, i.e., Dublin core and MARC. It incorporates a server that can serve any collection over the Open Archives Protocol for Metadata Harvesting (OAI-PMH) and it can harvest documents over OAI-PMH and include them in our collection. Export and import with DSpace is possible using Greenstone.

2.3 Interface

It has two different interactive interfaces: the Reader interface and the Librarian interface. End users can access the digital library through the Reader interface, which operates within a web browser. The Librarian interface is Java-based graphical user interface that makes it easy to download material for a collection from the web.

2.4 Metadata formats

It provides facility to users to define metadata interactively within the Librarian interface. These metadata sets are predefined: Dublin Core, RFC 1807, NZGLS (New Zealand Government Locator Service), AGLIS (Australian Government Locator Service). New metadata sets can also be defined using Greenstone's Metadata Set Editor.

2.5 Document formats

The system is extensible and customizable, i.e., plug-ins can accommodate different documents and metadata types according to the user's requirements. Plug-ins is used to ingest documents. For textual documents, there are plug-ins for PDF, XML, HTML, RTF, Plain text, Excel, PPT etc.

2.6 User base

It has a very strong user base, of these 80% are windows binaries and 15% are Linux binaries. The number of countries represented is above 70 and has more than 600 people on Greenstone email list. Examples of public Greenstone collections can be found at around 30 institutions and organizations including the California University, Chicago university library, Illinois Wesleyan University, Peking University Digital library, Vietnam National University. In India, Association of Indian Labour Historians, Delhi has developed digital library using Greenstone open source software.

3. DSpace

Dspace: Open Source Digital Library System is another popular software for developing digital libraries. It has been jointly developed by MIT (Massachusetts Institute of Technology) Libraries and Hawlett-Packard Lab. It is available under BSD open license for research institutions. DSpace is very good system for capturing, storing, indexing,

preserving, and redistributing digital research materials. It is most appropriate for Institutional Repositories to manage and serve the intellectual output of an institution. Universities and research institutions worldwide use DSpace as an institutional repository, a learning object repository, for records management, and more. Dspace manages and distributes the digital items and allows creation, indexing, and searching of associated metadata to locate and retrieve them. Dspace is a digital asset management system. DSpace open source platform is freely available so that one can customize and extend it to adapt to various community needs. It helps in creating, indexing, and retrieving different forms of digital items. Interoperability between systems is built in and it adheres to international standard for metadata. Dspace accepts varied forms of digital materials, like:

- Theses/dissertations
- Articles, preprints, technical papers
- Published books
- Administrative records
- Computer programs
- Multimedia publications
- Images
- Audio files
- Video files
- Web pages, etc.

3.1 Technical support

It has very good technical support. Dspace developers share information, help each other and solve technical problems through dspace-tech. Developers use the “Dspace Wiki” to keep in touch with one another, and to share and contribute code fixes and enhancements. DSpace runs on any UNIX or LINUX operating system. It can be easily installed and maintained and provides Dspace –tech, where we can meet other DSpace developers.

3.2 Document formats

Plug-ins is used to ingest documents. For textual documents, there are plug-ins for PDF, XML, HTML, RTF, Plain text, Excel, PPT, etc. Dspace supports different types of file formats including text, audio, video and images.

3.3 OAI Support

DSpace is OAI-PMH compliant. It defines the metadata standards based on Dublin core for the purpose of interoperability. It uses OAI-PMH through OAICat (an open source

product of OCLC) tool for harvesting metadata, which can be easily extendable to other metadata schemes by developing java programs. Moreover Dspace by default uses qualified Dublin core set (65 elements) for furnishing metadata and exposes metadata using unqualified Dublin core (15 elements) format for the purpose of OAI-PMH. The recent versions allow clientele to define their own metadata formats by using XML input forms.

3.4 Institutional Repository (IR)

DSpace is well established as an institutional repository too, i.e., digital library software in which scholarly publications in different forms (audio, video, images, etc.) are archived and preserved. It is found more appropriate to capture, store, preserve and redistribute the research output of an institution. It is known for its ability to accommodate the multidisciplinary and organizational needs of a large institution/university; providing access to the digital work of the whole institution through one interface and may organize into communities and collection, each retaining its identity within the repository.

3.5 Digital Preservation

DSpace provides long term physical storage and management of digital items in a multi-layered security, professionally managed repository including standard operating procedures such as back up, mirroring, refreshing media and disaster recovery. It identifies two levels of digital preservation. Of these one is Bit preservation which ensures that a file remains exactly the same over time- not a single bit is changed-while physical media evolves around it. Functional preservation goes further, the file does change over time so that the material continues to be immediately useable in the same way it was originally, while the digital formats evolves over time. DSpace assigns a persistent identifier to each contributed item to ensure that it is retrievable far into the future.

3.6 Features of DSpace

- Conforms to the standards like Dublin core and OAI-PMH.
 - Digital creation, management and distribution.
 - Security can be built at various levels to affect restricted access.
 - Submission facility allows librarians and researchers to upload digital items from anywhere in the world.
 - Digital preservation.
 - Supports uploading and downloading of bit streams as it is. Response is good with regards to commonly used file formats such as PDF, MS Word document, spreadsheets.
 - It provides long term physical storage and management of digital items.
 - Indian languages based digital libraries can be built as it conforms to the UNICODE standard.
-

4. Comparative Analysis

Comparison between Greenstone and DSpace has been carried out on points like the web address, availability, version, developer, prestige, operating system, system requirement, license, language, technical and training support, examples, security, browse and search which is as follows.:

Feature	DSpace	Greenstone
URL	http://www.dspace.org/	http://www.greenstone.org/cgi-bin/library
Availability software	Free open source software Download at http://sourceforge.net/project/dspace	Free multilingual open source Download at http://prdoenloads.sourceforge.net/greenstone
Latest version Developer	1.3.2 (5th Oct, 2005) Massachusetts Institute of Technology (MIT) Libraries and Hewlett-Packard Company	V2.71 (11th Oct, 2006) New Zealand Digital Library Project, University of Waikato, UNESCO, Human Info NGO
Reputation more reputation	Reputed and known in	Association with UNESCO gives the field reliability and adds to
Platform	Linux Or Unix	Windows (95, 98, NT, 2000, XP), Linux, Mac OS X, Unix
System requirement language.	Apache Web server, Tomcat Servlet engine and Postgre SQL relational database, Java 1.4 or later	Apache Web server, C++ complier and My SQL database, Perl Apache Web Server may not be required downloading on PC for testing etc.
Language used	Java	C++, Perl, Java
License	The BSD License	GNU General Public License
Language languages Hindi,	English	English, French, Spanish, Russian and Kazakhs. The reader's interface is available in many other including Indian languages: Bengali, etc.
Technical support	Technical mailing list for Discussions: http://www.dspace.org/feed back/ mailing.html technical questions.	Online support: http://www.greenstone.org/cgi-bin/ library?e=p-en-home-tfZz8&a=p&p= support

Training support worldwide instructional	Training support is not good as with Greenstone	UNESCO sponsored training programmes are organised including India. Self module available at http://www.imarkgroup.org/
Use Examples	MIT http://dspace.mit.edu/ ETD of IISc, Bangalore http://etd.ncsi.iisc.ernet.in/ Digital Archives for Indian Universities of INFLIBNET Ahmedabad http://dspace.inflibnet.ac.in/	New Zealand Digital Library Project, http://www.nzdl.org IIT, Mumbai http://www.library.iitb.ac/ Association of Indian Labour Historians, Delhi National Centre for Science Information, Bangalore
Security in DSpace.	Known for foolproof security. system.	Security system is good but not as It has multi layers security
Browse & Search	Supports search & browse, field based search and browsing menu driven. It is easy to use	Supports field based search, full text indexing, use. search & browse documents formats like HTML, word, PDF, PS; Unicode based multilingual support for user interface & content. It is quite easy

5. Conclusion

Importance of Open Source Software (OSS) in creating their own digital library has been realized by LIS professionals. It provides a new way of organizing and publishing the information on Internet or on CD-ROM. In this commercial world OSS has gained the popularity among information community as these are freely available to use and create their own digital library collections. End user is free to make changes and modifications before redistribution. The comparative table given above will help the librarians and information specialists to gauge the similarities and differences between Greenstone and DSpace. It can be concluded that DSpace and Greenstone both can be used for capturing, storing, indexing, preserving and redistributing the content. Both have been tested with creation and maintenance of a wide variety of digital collection all over the world. A major feature and plus point with both of these software has been the involvement of digital library researchers in their development. Some features of Greenstone like its multilingual nature, exhaustive platform and training program supported by UNESCO, have proved to be superior over DSpace. It is believed that with the support and promotion of UNESCO, Greenstone digital library software will soon become as popular as CDS/ISIS. DSpace is found more appropriate for building institutional repositories. It is functionally richer than Greenstone and supports a wide range of items including text, audio, video, and image. It provides a detailed

implementation guide. In India DSpace has been tested in building digital libraries, particularly the institutional repositories. There are two outcomes of the use of DSpace in India - Librarians Digital Library of DRTC(Bangalore) and Digital Archives of Indian Universities, INFLIBNET (Ahmedabad).

References

1. Naik, Umesha & Shivalingaiah, D. Digital library open source software: a comparative study. In proceeding of the CALIBER-2006, 4th International convention organized by INFLIBNET Centre in collaboration with Gulbarga University during Feb. 2006.
2. Nikam, Kaiser and Jayanthi, N. Digital Library Open Source Software: A Comparative analysis. In proceedings of the National Conference of Information Management in Digital Libraries (NCIMDiL) 2006 organized by IIT Kharagpur during August 2006
3. <http://www.opensource.org> visited Nov. 2006
4. <http://www.oss4lib.org> ,, ,,
5. <http://www.dspace.org/> ,, ,,
6. <http://www.eprints.org> ,, ,,
7. <http://www.fedora.info/documentation> ,, ,,
8. <http://www.greenstone.org/cgi-bin/library> ,,
9. <http://www.infolibrarian.com/dlib.htm> ,, ,,
10. <http://www.sourceforge.net/projects/greenstone> ,,

BIOGRAPHY OF AUTHORS



Dr S N Singh is presently working as Reader in the DLIS, Mizoram Central university, Aizawl - 796001. Before joining the Mizoram University he has worked as Sr. Assistant Librarian in the Central Institute of Higher Tibetan Studies, Sarnath, Varanasi. He has rich profession and teaching experience of about 18 years. He has around 20 articles/research papers published in various reputed journals and proceedings of seminars/conferences. He has also written two books.

Email: drsinghsn@yahoo.com

Mr R K Ngurtinkhuma, Senior Lecturer at Department of Library and Information Science, Mizoram University, Aizawl -796001.

Email: rkn05@rediffmail.com

Mr Pankaj Kumar Singh is Researcher (JRF) in Department of Library & Information Science, BHU, Varanasi -221005. He has 3 papers to his credit.