**A Comparative analysis of Digital Library Open Source Softwares:‘Dspace’ and ‘GreenStone’**

Lakshmi Verma

DESIDOC

Email id-laxmi@desidoc.drdo.in

**ABSTRACT**

The exponential growth in data generation and subsequent transformation into knowledge has created huge repositories of knowledge in the libraries. This has revolutionalized the methods and techniques to retrieve the relevant and useful information for the users. The growth of Information and communication technology (ICT) has facilitated into achieving this. In this paper, a study of two open-source digital library management software has been presented which collects and disseminates information for library-users. This analysis involves the study of related software documents and related technical manuals.

**Key Words**

Open Source, Digital Library, Digital Library Management System, software, DSpace, GSDL, Greenstone

**1. Introduction**A place, where collection of information resources is stored in print and other forms in an organized and accessible manner for print or study is referred to as Library. As defined by International Organization for Standardization a library is “irrespective of the title, any organized collection of printed books and periodicals or of any other graphic or audio-visual materials, and the services of a staff to provide and facilitate the use of such materials as are required to meet the informational, research, educational or recreational needs of its users”1.

Digital Library is a type of information retrieval system where the information is stored in digital format which can be accessed within network of computer users6. It uses online repositories which can store the textual information systematically and can be accessed by users 24X7. There are various such digital repositories available, which may be open source or proprietary.

Open source describes the method of software development, which uses the power of review and transparency of distributed peer-to-peer progression. Here the codes of software are available in open domain which can be customized by the respective users. This technique helps in providing high quality software through high reliablity, low cost, flexibility and end of traditional seller lock-in.Since, these open source software come under “Open Source free license”, it allows the developers / users to change, improve and distribute software many times.

**2. Digital Library Management Systems**

 The Open Source Digital Library management systems are the software applications that help in creating and presenting information repositories. The repositories built with the help of these Digital library management systems can be searched and browsed based on Metadata as these features are inbuilt in such applications. Apart from this, they can be easily maintained, enhanced and re-created. Presently many open source software (OSS) applications are available for library and information management. For example DSpace, GSDL, Fedora, E-print etc. Therefore, organizations can choose the one which is the most suitable for their requirement and implement them to create digital repositories. In this paper will be focusing mainly on two of the most popular Open source Digital Library Management Systems: DSpace and GSDL.

**3. DSpace**

 [DSpace](http://www.dspace.org/index.php) is an [open source](http://en.wikipedia.org/wiki/Open-source_software) library management software which allows us to capture and store digital data like text, video, audio etc into created repositories. It also provides facility to index, preserve and disseminate the digital material.  Thus digital libraries use DSpace to manage the digital materials and publications in professionally maintained repositories.

 If we see the world-wide scenario, there are more than[1000 digital repositories](http://www.dspace.org/index.php?option=com_formdashboard&Itemid=151&lang=en) which are developed using the DSpace application for storing, distributing and preserving their digital data.  DSpace is more common as a platform to build an [institutional repository](http://en.wikipedia.org/wiki/Institutional_repository)  which is a digital collection of research documentation, intellectual publications, library collections etc.

 DSpace performs three major tasks to build a repository:

* It captures and ingests the digital content along with metadata
* It lists the content systematically and helps in searching based on keywords and metadata
* It supports preservation of the digital data for a long period of time

 Therefore, DSpace can easily be customized to manage and preserve the digital content and provide accessibility of this data to the users. Since it is an open source software, an active community of developers, researchers and users across the world are collaborating to provide their expertise to enhance this application.

 DSpace is capable of storing a wide range of digital data, which includes documents like articles, technical reports, conference papers, books, theses, multimedia publications, Administrative records, images, audio-video files, web pages etc. It also provides multiple features like visualization, simulation of the stored data etc.

**3.1 Latest Features of Dspace**

As DSpace is a continuously growing platform, it keeps on releasing upgraded versions from time to time. 6.x is the latest update to the DSpace platform8. It consists of an upgraded configuration system, upgraded file storage plugins, and better quality control / health-check reporting features (through REST API and also through email).  Furthermore, DSpace 6 has a Java API refactor that adds support for both UUIDs and Hibernate in the database layer. This feature makes it compatible for future challenges.

As reported by DSpace official website, the new Features and improvements in 6.x version includes:

* “Java API refactor, featuring Hibernate and UUIDs
* Enhanced (reloadable) configuration system, featuring a new local.cfg configuration file
* Enhanced file storage plugins, featuring support for Amazon S3
* Configurable site healthchecks via email
* XMLUI framework for metadata import from external sources, featuring support for PubMed imports
* XMLUI export of search results to CSV (for batch editing)
* XMLUI extensible administrative control panel
* REST API Quality Control Reports, along with sample HTML clients and CSV export (for batch editing)
* REST API support for additional authentication methods (e.g. LDAP, etc)
* All searches default to boolean AND.
* Enhanced indexing for searches (Excel is now searchable, as well as right-to-left text in PDFs)
* OAI-PMH adds compliance for Open AIRE 3.0 guidelines for literature repositories”2

**4. GSDL**

 Greenstone Digital Library is an open source, multilingual software, which has been released under the terms of the GNU General Public License and is used widely for creating repositories and making them accessible online. The development and distribution of GSDL is an outcome of the joint efforts bythe [New Zealand Digital Library Project](http://www.nzdl.org/) at the [University of Waikato](http://www.waikato.ac.nz/), [UNESCO](http://www.unesco.org/) and the [Human Info hyperlink "http://humaninfo.org/" NGO](Human%20Info%20%20hyperlink%20%22http%3A//humaninfo.org/%22%20NGO).The aim of Greenstone software is to enable the users in building their own digital libraries. It provides a way to organize this information and publish it on the web or any other digital storage media like DVD and USB flash drives. In the later case, it will run on a non-networked environment. The digital libraries built by GSDL are fully-searchable and metadata-driven digital resource9.

 Infact, this software encourages the effective deployment of digital libraries to share information and put it in the public domain. Therefore, it is in itself not a digital library, rather it provides a platform to build the digital library.

 In 2004 its developers of GSDL were awarded by IFIP Namur award for "contributions to the awareness of social implications of information technology, and the need for a holistic approach in the use of information technology that takes account of social implications”9.

**4.1 GSDL Versions:**

There are two main versions of GSDL namely GSDL2 and GSDL3. GSDL2 was the earlier version and still under wide-use where as GSDL3 is the latest version under active development. The best thing is that GSDL3 has backward compatibility and contains almost all the features of GSDL2. If a programmer is already working on GSDL2, he can either work with the latest release of GSDL2 or consider upgrading to GSDL3. The Greenstone Librarian Interface (GLI) provides a feature to import 'Greenstone2 collection' which helps in migrating to the new software for existing users of GSDL2.

Greenstone3 has been developed in JAVA and uses various latest web technologies—like XML Transforms (XSLT), and the Java Authentication and Authorization Service (JASS). In the same context if we see Greenstone2, then it was written in C++ and was based on many self-developed techniques by the developers as many latest web technologies were not available at the time. This made the users totally dependent upon the documentation by the development team. All these limitations have been overcome in the latest GSDL version.

**5. Comparison of DSpace and GSDL:**

Based on above discussion Features Comparison Table for DSpace and GSDL is given below:

|  |  |  |  |
| --- | --- | --- | --- |
| **S. No** | **Features of Open source Software** | **GSDL** | **DSpace** |
|  | Product Type | Software | Software |
|  | Year of creation | 1997 | 2002 |
|  | License cost/Update Cost | Free | Free |
|  | License | GNU | GNU |
|  | Services | Training | Service via 3rd part service provider |
|  | Plug-in extends | Yes | Yes |
|  | Resource Identifier | No | CNRI Handles |
|  | OAI-PMH | Yes | Yes |
|  | Z39.50 Support | Yes | No |
|  | Supported File formats 7 | doc, pdf, html,ppt, postscript, jpeg, gif, video, mp3, etc | doc, pdf, html, ppt, jpeg, gif, audio, video,etc. |
|  | Supported Item Types(Storage and rendition) | Can store and manage all types of content | Can store and manage all types of content |
|  | Thumbnail Preview  | Images, Audio, Video | Images  |
|  | Multilingual Support5 | Greenstone provides ready-to use multilingual interfaces that are already translated in many languages. | Unicode character encoding, so different languages can be supported |
|  | Machine-to Machine Interoperability.  | Z39.50, OAIMHP | OAIMHP,OAIORE, SWORD, SWAP |
|  | Syndication | --- | RSS, ATOM |
|  | User Authentication  | User Groups | LDAP Authenticat ion, Shiboleth Authentication |
|  | Searching Capabilities4 | Field Specific, Boolean Logic, | Field Specific, Boolean Logic, Sorting options |
|  | Browsing Options | Browsing can be done | By Author, Title, Subject and collection |
|  | Metadata formats3 | Dublin Core, Qualified DC , ,METS, RFC1807’ NZGLS (New Zealand Government Locator Service), AGLS (Australian Government Locator Service) | Dublin Core, Qualified DC, METS |
|  | Associated Software3 | Apache Web server, Java 1.4.0 or above, Image Magick Software Ghost scripts and Web Browser | Java JDK5 or later Apache Ant 1.6.2 or later, Apache Maven 2.0.8 or later Java 1.4 or later, PostgreSQL 7.3 or later, Apache Tomcat 4.x/5.x and Web Browser |
|  | Software Platforms3 | Windows95/98/Me/NT/2000/XP/10 Unix/Linux, and MAC OS-X | Windows(NT/2000/XP/10) and AllPOSIX (Linux/BSD/UNIX-like OSs),OSX |
|  | Statistical reporting  | Yes | Yes |
|  | Databases  | Its Own | PostgreSQL |
|  | Programming Language  | C++, Perl, Java | Java and JSP |
|  | Web Server  | Apache/I IS | Apache and Tomcat |
|  | URL for free downloads10 | <http://www.greenstone.org/download> | <http://www.dspace.org/> |

**6. Practical Implementation Of Dspace At DESIDOC**

 DESIDOC is the information centre of DRDO which provides information to various DRDO laboratories through its e-services like Institutional repositories of DRDO, Archiving of Newspaper Clipping, Union catalogue of periodicals, Archiving of E- journals etc. For this, it uses library management system to create and disseminate information to the users through digital repositories. The Library management system used at DESIDOC is DSPACE as DSpace provides many inbuilt features like full text search, metadata based search, federated search etc. At the same time many more features can be incorporated in this DSpace digital repositories like usage pattern analysis, implementation of business-intelligence tools etc, to make the services much more effective and user- friendly.

**7. Conclusion**

 Therefore, Digital Library Management software (DLMS) provide a platform to the digital library-service providers to create an easy to use, customizable architecture for its users. With help of these, the institutional repositories, research documents, manuscripts, audio-video data of organizations can be stored, preserved and also disseminated to the targeted users. The two software discussed above, though differ in their architecture and presentation, still meet all the broad requirements of digital libraries. As a result, its difficult to prefer one specific DLMS over the other system. Instead of generalizing the suitability, we should emphasise on specific needs of a particular digital library. Depending upon the specific needs one DLMS may be preferred over the other. Theselection of the softwareis normally based upon on the format of data to be uploaded, the way it is to be disseminated,the choice of backend and frontend of the application and time duration for establishing a Digital Library etc.

**8. References:**

1.[http://portal.unesco.org/en/ev.phpurl\_id=13086&url\_do=do\_topic&url\_section=201.html](http://portal.unesco.org/en/ev.phpurl_id%3D13086%26url_do%3Ddo_topic%26url_section%3D201.html)(accessed on 6Dec 2017)

2. DSpace 6.X System Documentation . <http://www.dspace.org/>. (accessed on 6Dec 2017)

3.https://arizona.openrepository.com(accessed on 6Dec 2017)

4. Shahkar Tramboo, Humma, (2012), Journal of Computer Applications (0975 – 8887),59(16) retrieved from <https://arxiv.org/ftp/arxiv/papers/1212/1212.4935.pdf>

5. <https://researchtrend.net> (accessed on 6Dec 2017)

6.presentation on digital library <https://www.slideshare.net/sandeepsinghsainimba/digital-library-15882846> (accessed on 6Dec 2017)

7.Ravikumar and Ramanan( 2014), Journal of the University Librarians Association of Sri Lanka, 18(2), <https://jula.sljol.info/articles/10.4038/jula.v18i2.7867/galley/5202/download/>(accessed on 6Dec 2017)

8. DSpace 6.X System Documentation retrieved from[https://wiki.duraspace.org/display/DSDOC6x/DSpace+6.x+Documentation](https://wiki.duraspace.org/display/DSDOC6x/DSpace%2B6.x%2BDocumentation)(accessed on 6Dec 2017)

9.<http://www.greenstone.org> (accessed on 6Dec 2017)

10. Shalini R. Lihitkar\* and Ramdas S. Lihitkar,2012, DESIDOC Journal of Library & Information Technology,32(5),393-400 retreived from <http://publications.drdo.gov.in/ojs/index.php/djlit/article/view/2660/1318>(accessed on 6Dec 2017)