Chapter 13

Conclusion

13.1 Summary

Open source software (OSS) have gained lot of acceptance in libraries all over the world to save exorbitant costs incurred in procuring commercial software as well as paying annual maintenance costs and upgrading costs incurred by commercial suppliers.

A range of open source software are available today for use in libraries for various applications. Use of open source software in libraries has made a good impact in libraries because library profession is highly collaborative and has a history of sharing successful approaches.

UNESCO has provided CDS-ISIS free to libraries for automating bibliographic databases. With the help of CDS-ISIS libraries could manage to automate their bibliographic databases where budgets were the major obstacle for the libraries. However though CDS-ISIS was distributed at no cost it was not an open source software, it was just freeware. The end-user can only download binary files for MS-windows platforms for using CDS-ISIS. The approach of not providing the source code of CDS-ISIS restricted its use only to MS-Windows and hardly made usable on UNIX/Linux or Solaris platforms. Indeed, SCO-UNIX and VAX/VMS binaries were distributed for some time with no GUI support.

Today, quite a few Open Source Software (OSS) are available on Internet especially for creating digital libraries/digital repositories/institutional repositories/digital archives. There are many other applications as well available through open source license terms and conditions, such as Library Management Systems, MARC modules, creating citation programs,
barcode generators, Z39.50/server protocol, federated search engines that can search across electronic journals, OPACS etc. Libraries across the world have been very successful in making use of these OSS tools and sharing data globally. OSS tools can be effectively used to successfully bring libraries important data on web by participating in the Universal Access to Information and giving more visibility to their own users.

The success of any open source project depends on the clarity of the shared vision of the goals of the software and some strong definitions of basic functions such as how they work. Many of the OSS application developers have proper funding support from different funding agencies for sustainability due to good user base. The developers of these software are continuously working to make these software more user friendly with advanced features. In the development of any OSS many people play different roles in making OSS successful. One need not be a programmer to contribute, one can help in preparing documentation manuals, reporting bugs, presenting wish-list, evaluating the software or usability testing.

From the late 1990s, there has been tremendous emphasis on digital library, mainly as a scholarly publication tool and to achieve the bigger goals of open access to information movement. Initially, Greenstone software was made available to all under GNU Public License during 1997, for creating digital libraries/digital repositories. Along with Greenstone, there are many other repository software available today under Open Source License terms and conditions. These software are equipped with DL functionalities such as storage, search and retrieval of digital documents along with a user interface.

During 2000, open access movement came in scientific publishing and Steven Harnad argued that every author should have access to his/her paper irrespective of where it is published and copyright should be withheld with the authors. Due to the open access movement, many authors started building preprint/postprint collections of their papers either through their personal web sites or through institutional web sites. EPrints software was made available free under the open source license during the same time and many organizations as well as authors started making use of EPrints software for creating web based archives of scholarly output of authors. By making preprints/postprints available free on Internet, author’s got more visibility and sustainability.

Authors also started getting more citations to their papers due to availability of their papers openly on Internet. Slowly, more number of open source digital library software (OSS-
DL) evolved during the past decade for creating and maintaining institutional repositories/digital libraries and many libraries actively participated in using either EPrints, Greenstone, Fedora, DSpace, CDS-Invenio, DoKS, MyCORE, SOPS, etc.

Several libraries made an attempt to use one of these software either by their popularity or by knowing its functionality and good features supported by each software. But still majority of community lacked in knowing how many OSS-DL are available for creating digital libraries, what features do they support, how big is the user base of each software, what are hardware/software requirements for successful implementations of any of these software, whether software has digital preservation support etc. Considering these points, one of the aims of the present study was to first identify OSS-DL available on Internet for building digital libraries/institutional repositories. Then next objective was to define an extensive list of evaluation criteria for evaluating selected software taken for the study. The objective was further extended to installing each selected candidate software on a test bed environment to evaluate each criterion.

So far, very few studies have been carried especially in evaluation of open source digital library software. Chan(2004), OSI(2004) and Goh (2006) carried out preliminary evaluation studies but these studies lacked in-depth evaluation as well as they did not cover all OSS-DL software.

In general, we often come across many solutions for a given problem. However, we tend to choose a solution which best suits our needs in a given set of environment. This is also typically valid while selecting digital library software. Considering this fact, the present study does not attempt to rank the software and indicate the best, rather it attempts to present various criteria which can be defined while evaluating any OSS-DL and features offered by each software, so a library may choose the software which suits best to its user needs and in its given environment.

After a careful study of the literature published, an extensive list of evaluation criteria for evaluating OSS-DL has been developed.

During the course of study all the selected DL software were installed in Linux environment to create test-beds for each software, with the view to ascertain ease of installation and more so to test the claims of each software with regard to the features they offer.
A test bed environment was created on Linux Operating System with Fedora Core 7 version of Red Hat with the hardware of 1 GB RAM, 160 GB hard disk. An attempt was also made to install the stable version of each software where ever possible.

Though in the present study, more emphasis was given on successful installation of each software, an attempt was also made to build a sample collection of different types of documents along with different file formats by setting different work flow processes, setting authorization policies etc. The collections built in each software covers text, audio, video, image, files etc. A collection of users were also created in each software to ascertain how each software handles different policies for different users. This is necessary to understand functionality of each software and to find out different features supported by each software.

13.1.1 Benefits of using OSS-DL

One of the major advantages of using any OSS is no cost is involved. The OSS applications can be installed on different platforms such as MS-Windows, Linux, Mac OS X, Solaris etc. There is a community of developers/technical users/general users in each software which constantly exchange ideas through mailing lists hence there are more chances of getting support for any problem that we encounter while using any OSS.

The latest versions of the OSS are immediately available to the end users without any costs. Since source code is available end user can make changes as per his or her own requirements in the main program.

However, it should be noted that, not only the software should be open source, but also the software tools used in the software are open source. For example, ARNO digital library software is open source, unfortunately, it uses ORACLE as a back-end database. Precisely, for this reason the present study has not considered ARNO for further evaluation. In other words, the software should not have any hidden costs, it should be completely available free with access to source code.

13.1.2 Limitations of using OSS-DL

The trade off with open source software is, as the software in many cases is not distributed as binary files, one has to compile the software. Compiling using the respective language
compiler is more a programmer’s job, rather than a library professionals work. Besides compilation, making the software work in a web environment, one need to know the nuances of web server configuration. Additionally, it is required to have general system administration knowledge.

Similarly, most of open source digital library software are complex to install. Installation of essential components such as databases, network settings, web servers is often difficult to configure. In most cases, library professionals have to depend on in-house system administrators help to handle these jobs. Of course, there are few exceptional librarians who venture in learning and exploring Information Technology.

Many OSS mailing lists discussion forums show that considerable effort is required to successfully implement any OSS-DL as well as a dedicated team is required to help people solve their myriad problems which they encounter while using the OSS-DL. For using any OSS-DL since there is no specific vendor responsible for providing support, mailing lists are extensively used by the OSS user community to solve their problems and different mailing lists for different groups such as developers’ mailing list, technical mailing list and users’ mailing list are available for the users. These lists are very useful to get answers to simple queries as well as complex queries.

13.2 Learnings from the present study

While evaluating the OSS-DL, different models were observed. Each software has a different purpose. CDS-Invenio was initially meant for managing preprint server then later on it was developed for managing variety of digital documents along with managing bibliographic data of any library. DoKS was established to organise electronic thesis and dissertations as well as curricula vitae of graduate students for marketing themselves through DoKS repository. DSpace was developed for building and deploying a repository system to capture, preserve and disseminate an organisation’s born-digital assets to help faculty members to publish their intellectual output on the institute’s repository to get more visibility.

EPrints was created to facilitate authors self archiving their preprints/postprints via web and software got more acceptance among the scientific community as it is easy to self archive
using EPrints software. Fedora was developed to describe a digital object and repository architecture for storing and disseminating digital library content. Greenstone was created for building and distributing digital library collections. It is not a digital library but a tool for building digital libraries. MyCoRe was developed for organising multimedia teaching and learning collections. SOPS framework was designed for supporting publishing aid for conference organizers and organising digital collections. Thus these software have some common objective such with the case of EPrints and DSpace and each software addressed different facets, different audience.

The following sections briefly present a small summary of features supported by each software.

CDS-Invenio has been mainly defined for creating digital library/institutional repository along with managing bibliographic data of books and journals. It can act as a digital library/institutional repository of an organization and also as an OPAC of the library. It can be configured for different disciplines. It has very good workflow support and authorization policy support. It allows to define different roles for different users on different objects with help of defining complex authorization policies which is one of the important feature of CDS-Invenio.

Software allow to support hyperlinking multimedia digital objects. For example, a presentation by a faculty can have two links such as link to video file and a link to presentation file in PPT or ODP format, and also a link to the full text of the presentation in MS-WORD or ODT format. Software allow to link these different file types together with one hyperlink. CDS-Invenio also has unique feature which is not there in any other popular OSS-DL software such as write reviews, write a comments, add a particular document to individual’s personal collection basket, add citations etc. It also allow end users to rank each document added into the repository by the end-users. It allows to initiate discussion groups as well as supports personalization feature. Personalization is yet another active area of research, as it is expected to filter noise from information glut. CDS-Invenio takes care of many advanced features.

One of the strengths of CDS-Invenio is underlying MARC21 metadata format as basic format for data entry hence it is possible for the library to link metadata of the print copy along with the digital copy easily. CDS-Invenio can handle millions of digital files hence
CHAPTER 13. CONCLUSION

scalability is not an issue in CDS-Invenio.

Software automatically extracts words from the documents which helps to build a controlled vocabulary. The search engine of CDS-Invenio is very powerful which is Google like search engine and it can combine metadata based search and full text along with citation search. Software allows to export data in OAI-MARC 21 in XML and also supports OAI-PMH protocol.

One of the weaknesses of CDS-Invenio is its installation. For the successful installation of CDS-Invenio one should have a UNIX system administrator who can handle different installations that are required for CDS-Invenio. The software has many good features but the installation of the software is very difficult as well as the software has not yet reached even to version 1.0 though the first version i.e. 0.0.9 was made available during 2002. The development is very slow and due which there are less repositories using CDS-Invenio as well as there are no training programs conducted on using CDS-Invenio. Overall the software has been found a good choice for creating digital libraries/institutional repositories and organising OPACs.

DSpace jointly developed by MIT and HP Labs, is designed to operate as a centralized, institutional service. The DSpace model is best suited in the environments of universities and organizations where students and faculties bring out lot of scientific literature and which needs to be archived at central place as well as which needs more visibility.

After the formal release in 2002, DSpace got wider acceptance from all over the world and there are many installations of DSpace today. DSpace has been found as the largest used DL software among all the OSS-DL available today. Software supports adding all types of digital objects from audio, video, text, image to datasets and supports Dublin Core as a metadata standard for data entry. It is considered as a trusted digital content repository system from digital preservation point of view.

DSpace helps organizations to provide different services such as long term access to faculties intellectual output, more visibility to faculty members by uploading their contents on the institutional repository thereby making available a parallel publishing channel, providing different authorization policies and support of persistent identifier i.e. Unique accession number to each document that gets added into the repository.
DSpace supports weed out policy for all documents that are added into the DSpace repository. The current release of DSpace also supports 'controlled vocabulary' feature which is also supported by CDS-Invenio, EPrints & MyCoRe.

DSpace is not very easy to install. One needs to have system administrative support for successful implementation and customization of DSpace. It works very well on any UNIX like OS such as Linux, HP/UX etc. operating system. There are couple of training programs held all over the world on DSpace. In India since 2004, DRTC has been sharing their experiences on using DSpace and by conducting hands-on training courses and helping on installations by running a mailing list called as dlrg@drtc.isisbang.ac.in

DSpace yet does not support digital object versioning as well as yet does not support any authority files control. DSpace only creates email ID and name of the user who has registered into the repository it does not support yet keeping other details of each user such as affiliation, department/organization, registered date, last accessed etc. Software yet does not support to upload documents from existing URL. Duplicate checking of entries is not carried out in DSpace.

DoKS developed by Katholieke Hogeschool Kempen (KHK) is developed for organising electronic thesis and dissertations of university students along with organising their curriculum vitae. DoKS can be used as a recruitment model where every student can upload his CV, full text of thesis, full text of papers etc. One of the important feature of DoKS is it allow end user to decide to what level access should be given to others for every document that is added into the repository. Software supports to add any type of MIME type file format and support ETD-MS metadata support. One of the important feature of DoKS is, it allow students to advertise, annotate and communicate with others through DoKS portal by submitting their thesis and CV with embargo (time period).

Software supports metadata search plus full text search from thesis added into the repository. Software supports personalized services. Software supports OAI-PMH protocol for metadata harvesting. One of the distinct feature of DoKS is it supports couple of Web 2.0 technologies such as social bookmarking, instant messaging, Google adsense etc. which are not supported by other software. Software also supports RSS and support storing queries submitted to DoKS. Software supports UTF-8 hence it is possible to create repository in any language.
DoKS does not have any digital preservation support. No new versions of the software are made available after 2006 hence it is difficult that software will remain available in future. The functioning of DoKS is as well not very easy to understand. One has to spend lot of time in understanding how DoKS functions. There is a very small user base of DoKS and there are no training programs conducted for using DoKS. There are no mailing lists available for the software. Hence, though software has some unique features it has not got wider user base.

EPrints developed by University of Southampton, UK has gained wider acceptance since availability of its first version during 2000. EPrints was basically meant for authors to archive their preprints or postprints to gain more access and visibility to their work. There are large number of universities which are using EPrints as a platform for archiving preprints /postprints.

EPrints version 3 has many more added features. The new version of EPrints software supports high metadata quality as well as it has very good data entry template with many added fields such as for thesis data entry it has title, author, abstract, thesis type (Masters/PhD/Eng/Other), Creators, funders, projects, Departments, Institutions etc. Hence for different types of documents EPrints has customized metadata templates. Among all the evaluated software EPrints was found best for metadata entry. EPrints also allows to create bibliographic database of documents and has the auto complete feature which is available for author, journal and ISSN field.

Software supports audit trail of all documents that are added into the repository means EPrints tracks all changes and actions to all documents that are deposited into the repository. No other software has this feature support and this feature will be helpful in future releases of EPrints from digital preservation point of view.

Software supports to add any type of digital document into the repository. While submitting documents software automatically recognizes the uploaded file format if supported by EPrints.

EPrints keeps detailed information of each user who has registered into the repository such as name, affiliation, address, email address etc. The work flow in EPrints at present is restricted to user, editor and administrator. It does not have any other work flow management.
EPrints supports weed out policy. It allows to put embargo date for all the documents that are added into the EPrints repository. EPrints also supports to add different versions of the files attached to one metadata set. Data added in EPrints is OAI 2.0 compliant. All documents added into EPrints repository are controlled by Library of Congress subject heading list.

One of the drawbacks of EPrints is it does not yet support any persistent identifier numbers for all the documents that are added into the eprints repository and has limited submission support. It also has limited workflow management. Access control in EPrints is controlled by user name and password. It does not have any other means to restrict access to EPrints. Digital preservation is not yet implemented in EPrints.

FEDORA is an open source digital asset management system. One of the main features of Fedora is that it allows end users to build a collection of digital objects through local servers or through http servers or it also supports to redirect the Fedora to look for a particular object from another place which is available online and give link to that object through Fedora repository. It is not necessary to have that object on Fedora’s site.

It is a modular system that meets the necessary functional requirements of a long-term repository of digital objects. Fedora uses Universal Resource Naming service which is called as URN for all the objects that are added into the repository as well as it allow end users to add other persistent identifiers as well.

Fedora supports comprehensive versioning framework for all digital documents that are added into the repository. It supports to track the different versions of the documents and supports to provides access to earlier versions. Version support is an essential feature in digital preservation, as preservation of earlier versions of digital document is essential for accessing all versions whether older or the latest and Fedora supports to upload different versions of digital documents.

Fedora is basically designed for handling millions of documents with long term preservation of documents added into the repository. It supports to upload complex and heterogeneous digital objects as well as big data files. It can handle any type of MIME file formats. Similar to that of CDS-Invenio & EPrints, Fedora allows to add bibliographic details as well as full text documents.
One of the distinctive features of Fedora is that it allows to see on the screen, different behaviors of each digital object. For example, an image can be viewed in Fedora in different scales. In other words, Fedora model supports to define different behaviors for all the digital documents that are added into the repository for e.g. Fedora can define different descriptions for an image to be viewed on web page such as getThumbnail, getImage, getResized, getZoomedImage, getImageTitle, getXML, getTOC, getDocument, getFile etc.

One of the weaknesses of Fedora is basic fedora installation does not support any workflow or authorization policies or added features. One has to install other tools which can be integrated with Fedora such as Fez, Valet, Murador, Elated etc. to have more added features. Fedora user interface is not very easy to use as well as it is not a web based software. FedoraAdmin user only can create collections in Fedora which is a similar function support in Greenstone.

During the evaluation of Fedora attempt has been made to integrate Fez with Fedora but the latest Fedora version does not support Fez, as the Fez latest file has bugs hence it was not possible to explore Fez with Fedora. Basic Fedora installation does not support advance search or other search features and requires to configure other fedora search tool which is called as FedoraGesearch. One basic disadvantage with add-on is to keep these add-ons up to date with the newer releases of the software and Fedora and its add-ons are not updated properly.

Basic Fedora installation is easy to carry out but one of the drawbacks of Fedora is it does not have any ready made metadata templates. End users have to type their metadata in XML form. Fedora supports to large extent digital preservation support which is one of the strong feature of Fedora.

There are no training programs conducted on Fedora. Fedora user community is very small as well as user interface of Fedora is not user friendly hence software has not got more acceptability.

Greenstone is one of the first OSS-DL solutions made available under Open Source License. It has made its impact all over the world especially in developing countries for building DL.
CHAPTER 13. CONCLUSION

One of the attractions of Greenstone is its installation. The installation of Greenstone is very easy and it works easily on any version of Windows, Unix and Mac OS-X. Librarian can easily install Greenstone without having any administrative support. Greenstone also has one more advantage as it helps to export full software along with built digital collections onto a CD/DVD. This feature is useful for the developing countries, who do not have good Internet connectivity. There have been several workshops, tutorials and courses conducted in the recent past. In India IIM, Kozikode has been identified as a ‘Greenstone support for south Asia’.

For a small library where budget and staff strength is less and if the library wishes to create a digital library of important documents available with them it is easy for that library to build their collections with Greenstone. Greenstone can be configured in variety of disciplines with variety of digital document types. It has the capability to handle different file formats but at the same time it does not have any capability to handle digital preservation part.

Greenstone can handle building collection of digital books. It has a very good feature of browsing table of contents, browsing through different sections of the book as well as it provides a very good search feature. If a particular library has a major collections of digital books, Greenstone can be considered as one of the options for organizing these books and making these collections available to the end user through Intranet/Internet.

Greenstone supports to have a variety of predefined metadata schema sets. A library can select metadata schema as per their requirements. Greenstone allows to upload digital documents from the local servers as well as directly from the web. It also has OAI-PMH support as well as Z39.50 server support which helps to import data from OAI-PMH servers or Z39.50 servers into Greenstone server directly. Greenstone also supports to export metadata to METS or helps to import metadata from METS.

Another unique strength of Greenstone is its multilingual nature. The user interface of the Greenstone has been translated into more than 45 languages. The recent versions of Greenstone supports a DSpace plugin where DSpace data can easily be imported to Greenstone and Greenstone data can be easily exported to DSpace[1]. During 2000 Greenstone, brought out a new feature called ‘Collector’ and marketing it as “The end user building collections” that gave more acceptance to Greenstone in library community as well as other
general users. But this feature is not very extensively used by the libraries. Libraries prefer building collections through Librarian’s Interface which is provided with Greenstone.

One of the limitations of Greenstone is that it does not support any work flow neither it allows to set different authorization policies. For every document that gets added into Greenstone does not have unique persistent identification number. In other words, Greenstone’s, model is not based on self-archiving.

One more serious disadvantage with GSDL is that one has to build the collection and generate indexes every time some documents are uploaded. This is due to the MG-II search engine which GSDL uses. Unfortunately, MG-II search engine does not have incremental indexing. However, in the recent versions of GSDL attempts are being made to switch over to Lucene Search engine, as it allows incremental indexing, and GSDL is being/will be offered using JAVA programming rather than the Perl of earlier versions. However, JAVA / Lucene based GSDL version 3 is not still a very stable product and there are many limitations in using successfully GSDL version 3.

MyCoRe supports to generate digital libraries and content repositories. The documentation and other manuals of MyCoRe are in German. Though limited support in English was available, the software has been found as one of the good software for organising digital collections. Software supports to add variety of digital documents with variety of file formats. Software supports URN for every document. Software has proper work flow where it can handle different submission rules. MyCoRe by default supports qualified DC schema.

One of the unique feature of MyCoRe is it identifies the file format and then only uploads those files into the repository. Unknown file formats are marked as ‘unknown’ and are uploaded into the repository. Software supports DDC classification. It also supports hierarchical classification of an organizational structure. MyCoRe has a very good information search and retrieval support as well as access control support. It supports searching with variety of search strategies.

One more unique feature of MyCoRe is it records complete details of every user that registers into the repository. MyCoRe has very good access control mechanism. Software allows to control access by individual user from the repository as well as allow to control access by IP authentication at different levels such as permissions can be set for every
group, every user for a particular period with read, modify and create options. Software supports Z39.50 protocol as well as OAI-PMH protocol. MyCoRe has not implemented any digital preservation support yet.

There are no training programs conducted for MyCoRe all over the world. The MyCoRe community at present is basically from Germany. The mailing lists at present are used by German community hence the language used for any communication is German. Hence MyCoRe has not got wider acceptability all over the world yet though it is a good software for creating digital repositories.

SOPS stands for SciX Open Publishing Services. SOPS is not very popular software. It has not brought out any further versions after its first release. Software supports digital library services with limited support such as it supports to add only electronic publications. Installation of SOPS is not very easy. Software supports to export data into variety of formats. Software does not have any work flow management for uploading documents though it is mentioned in the documentation manual. The overall performance of the software was not satisfactory. Software does not support many features such as authentication and authorization, classification, ontology support etc. It is just meant for organising digital objects with very limited support. Search and retrieval also is supported with limited features in SOPS. SOPS supports OAI-PMH protocol and it also supports to provide personalization feature such as software supports to initiate online discussion on a particular document submitted into the repository.

Software has not brought out any new versions, neither there are large number of users using this software. There is no mailing list available for this software hence it is difficult to predict the sustainability of SOPS in future.

### 13.3 Major Issues

#### 13.3.1 Metadata

Metadata is created and entered manually in the majority of the repository systems that were selected for evaluation. The future OSS-DL releases should have capabilities to extract technical metadata automatically from the digital objects which are loaded as well as
should have capabilities to support entering preservation metadata\[2\]. The current OSS-DL has limited metadata support. Fedora and CDS-Invenio has good metadata support but the other software still have limited metadata support. The existing software should be able to provide support for different metadata schema for different digital objects.

### 13.3.2 Persistent Identifiers

Only DSpace and Fedora & MyCoRe supports persistent identification number for every document that is added into the repository. Persistent identification is necessary from digital preservation point of view hence it is essential that software support persistent identification number for every document.

### 13.3.3 Scalability

Scalability is one of the important issues which needs to be considered by major OSS-DL developers. One of the hypotheses for the study was the OSS-DL are not scalable in terms of size. It was found that CDS-Invenio, DSpace, EPrints & Fedora can handle millions of records. There are proved results of scalability for these software. There are no proved results yet available for Greenstone, EPrints, DoKS, SOPS, or MyCoRe about handling millions of records.

### 13.3.4 Import/Export

Import/Export is one of the important features in any DL system. Each software will provide data export/import format into different metadata schema such as either Dublin Core/MARC/MARC XML/METS/FOXML etc. Current version of all software support metadata import/export utility. Actual data have to be separately copied from one software to other other for most of the software. An experiment was carried out of exporting DSpace based exported data to be imported into one of these software and it was found that none of the software were able to import DSpace data successfully. Hence it is very essential that all software should have one universal export data format such as XML format which can be used for importing into any software. The functionality of using import/export tools should equally be simple to use.
13.3.5 Upgrading Existing Systems

Upgrading from old version to new version is a continuous process and it should be easy to handle. While upgrading any software different factors needs to be considered such as back end database support, current OS support available etc. Hence it is necessary that all the software should have easy steps to upgrade their existing systems to newer versions.

13.3.6 Document Versioning Support

At present only EPrints, FEDORA, CDS-Invenio and MyCoRe support to add different versions of digital objects. This is one of the important feature that the repository software should support. Versioning controls change of format, 'reformatting' and change of content, 'revision' to the original document. It would be desirable for repository systems to provide standard and easy to use versioning management components for the long term preservation of digital objects. Additionally repository should support to provide, if one representation of the record is transferred from another representation, for e.g. transferring JPEG to TIFF, its derivative information and technical data should be recorded in the system in a way that retransferring can be performed when it is needed.

13.3.7 Sustainability

Sustainability is very important factor for any open source. Because open source projects depend on the interest and efforts of individuals it does not guarantee that they will continue in the indefinite future. Greenstone has a great difficulty to get funds as their university i.e. New Zealand University is not funding, neither UNESCO is ready to fund hence despite strenuous efforts, Greenstone is yet unable to develop a reliable sustainability model[3]. EPrints has been well funded by JISC, UK. DSpace has a support of MIT and HP. Fedora is also well supported by Cornell University, University of Virginia Library, Moore Foundation and Andrew W. Mellon Foundation. CDSware is well funded by CERN, Switzerland. MyCoRe is supported by Essen University and it will be supported in future. DoKS & SOPS have not brought out any new versions hence it is difficult to predict its sustainability in future.
The software should also support backward compatibility. At present none of the software support backward compatibility. This is an important feature and should be supported by all the software.

### 13.3.8 Standard Compliance

Interoperability among the existing repositories is very essential part to share metadata across institutions. All the selected OSS-DL software support OAI-PMH protocol for metadata harvesting. One of the hypothesis of the thesis was majority of the software are not conformity with the standards but it was found that all software support OAI-PMH protocol. Due to OAI support, users outside the institution are able to find and retrieve information from repositories and interoperability helps to create a global network of digital research materials.

### 13.3.9 Streaming server support

In future more of multimedia collections will be made available in DL hence it is necessary that the future OSS-DL should have capabilities to provide audio, video collections available through streaming servers. This will help end users to see large audio/video files instantly without any time delays. Many open server streaming servers are available today. Current OSS-DL should support to integrate with the available streaming servers. None of the present OSS-DL software have any mention of streaming server support feature.

### 13.3.10 Installation of the OSS-DL

Currently for installing any OSS-DL, expect Greenstone, users have to undergo a complicated manual process to install and configure the majority of the repository systems. There are several issues which users have to consider while installing these software such as code compiling, database and web server applications download, setup and configuration, customizing configuration files and many other things. The installation time largely depends on the experience of users. It is therefore necessary that the future OSS-DL should consider to provide easy installation steps for getting more acceptability from the librarian community.
13.3.11 User Friendly Interface

User demands will increase day by day and it is important that the DL software should provide more well-defined user interfaces which will help users to retrieve information more effectively without any delays. Interface is one of the most important aspects of usability as it is the medium that users communicate and interact with the system\cite{4}. The present versions of all OSS-DL provide very primitive user interface. It is necessary that future releases of these software should provide more user friendly user interfaces from user’s point of view.

13.3.12 Multilingual Support

To-date metadata schemas have been developed and applied in monolingual environments; end-user demands for accessing multilingual resources in the language of their choice will require new or significantly expanded and enhanced metadata schemas, or innovative applications of existing non-verbal metadata schemas to describe and retrieve multilingual resources. All existing software support UTF-8 i.e. Unicode Transformation Format which allow to create data in any language. Due to UTF-8 support it is also possible to create User Interface in any language. This is one of the important feature supported by all the software.

13.3.13 Copyright

The existing OSS-DL repositories should have capabilities to support keeping licensing details of all digital contents that are added into the repository. Copyright and license issues are very important in the digital environment. Today many small publishers are merging with big publishers and the online contents are made available to libraries with different license terms and conditions. Publishers are constantly changing license terms and conditions hence it is very essential that the selected OSS-DL should have capabilities to provide detailed licensing information about each digital document that is added into the repository. The software should have a proper feature which distinguishably shows license policies for every document that is added into the repository. Among the evaluated candidates DSpace supports distinctly entering license agreement for every document that
is added into the repository. EPrints allow to select type of license from the available list of license forms. MyCoRe also supports entering license details for every document that is added into the repository.

13.3.14 Digital Preservation

The digital repositories which are formed today are not well formed or structured yet. The number of repositories are growing internationally but these repositories are at different stages of development in terms of content, institutional backing, funding as well as in policy. One of the hypothesis of the thesis was digital preservation is not yet addressed by many of the OSS-DL. It was found that only Fedora software supports digital preservation to some extent. Such as all objects in Fedora are internally represented in XML form. It supports to check data integrity into different forms such as MD5, SHA-1, SHA-256, SHA-384, SHA-512, it supports to record different unique identification numbers for all documents, supports to keep all details of files such as different versions of files. Hence to some extent Fedora has capability to handle digital preservation.

DSpace at present only supports to store data into bitstream format. None of the other selected software have initiated any support with respect to digital preservation. Repository directory OpenDOAR discovered that only one-third of repositories available on Internet have any kind of digital preservation policy initiated[5]. Hitchcock(2007)[6] surveyed digital repositories and discovered that none had a formal preservation policy.

It is also very essential that the existing repositories make a policy for restrictions on accepting only open standard file formats for uploading into the repository. If the repository has the policy to restrict on open standards there are chances to successfully preserve document for several years without any data loss and migration of data from old format to new format would be easier and successful.

It is also necessary that all over the world research councils and other funding bodies need to mandate that all scholarly publications generated by publicly-funded research are made available on an open access basis then only the contents added into the digital libraries/digital repositories will grow. Library community needs to find ways to encourage academics to share and re-use publications, research data and learning resources as openly
as possible.

13.4 Future of Digital Libraries

In future more and more open repositories will be available on the Web from all over the world due to the advancements in hardware and software technologies and more so with gaining popularity of open access to information movement and use of open source software.

Digital libraries in future will function as an institutional entities with a key role to play in social, cultural and educational activities\cite{7}. In future, collaborative digital work environments, social digital navigation and new forms of digital environments for academic, educational, publishing and recreational purposes will be available for digital library researchers. More of Web 2.0 technologies will be needed to be used in future by the existing OSS-DL software.

The future of digital libraries will provide more personalised services such as personalized browsing using new information visualization and relevance feedback tools, novel interfaces, annotation of retrieved information. Providing putting annotations is an important feature of DL to support.

The next generation of digital library development and deployment will focus more on standardization, usability, and productization - providing greater usability for library patrons, increased interoperability among digital collections, and more cost-effective choices for institutions which are just beginning digitization programs.

The current digital libraries provide search using typed words yet searches based on spoken natural language queries are not supported by many of the OSS-DL due to the non-availability of speech recognition systems. Spoken natural language queries will be more useful while retrieving audio/video collections from the digital library. In any case this problem is in general applicable to any search paradigm not in specific only DLs.

The user interfaces supported by many of the current software are too primitive hence many advanced features are still not supported. The current digital library software yet do not support integrated searching or federated searching. Integrated searching such as retrieving
text, audio, video, image format i.e. Multimedia at one time will certainly be very useful to end users.

The future of digital library research hence will concentrate more on data delivery, user interfaces such as more advanced user interfaces, image understanding, natural language processing, speech recognition, federated searching etc.

Scalability as well is one of the important aspects which will have to be considered by the existing digital repositories to handle large amounts of data in varied formats. Software developers will have to consider how the software performs when there are more number of simultaneous users accessing the software at the same time. Each software should be able to handle millions of records with large file size.

Commercial systems such as DigiTools, CONTENTdm, Open Repository software etc. do provide similar digital library services to many organizations but they charge heavily and due to their commercial attitude there is little opportunity for advancement of any commercial solution. On the other hand, due to shared collaborative efforts in OSS there are many successful solutions of OSS available today.

The future will be with more reliable, distributed digital libraries with grid architecture. Future DL will support redundant storage and automatic data distribution, with approaches for ensuring content security in distributed environments and management of content and community work flows.

### 13.5 Suggestions

In general, overall OSS-DL developer community should make every effort to build the software in such a way that it is simple to install/compile which could be the important criteria for any OSS-DL to get the wider acceptance. During the present study all the selected software were installed on a single hard disk but there were many challenges in configuring each software on a single server due to clashing of same web server for two software, file permissions settings, disk space required for each software, consideration of security issues, creating different user communities, using different databases used by each software etc.
For the library community, it is very essential to have a simple installation of the OSS-DL otherwise in any organization if the software is difficult to install, it goes into IT department of that organization and then it becomes difficult for libraries to exercise their skills and demonstrate their values to the organizations and more to maintain the repository. Hence from the library community point of view, it is very essential to have simple installation process to the extent that anyone can install the software.

It is also essential that more number of training programs should be carried out at national, regional and local level on giving proper training on using each software. In this regard, Greenstone gets 100% marks as the installation of Greenstone is the simplest and it helps Librarian to build their collections very easily.

Generally the documentation manuals which includes user manual, developers manual, installation manual are difficult for end users to understand hence it is necessary that the OSS community should make an effort to create these manuals in more user friendly way to get more acceptability. Of course, there are WIKI pages available for popular OSS-DL which help them to get lot of information.

Any decision about which digital library/institutional repository/digital repository platform to choose must be based not only on the technical and functional capabilities of the system but also in determining best fit with organisational IT strategy and with the availability of local software development/administrative support as well as user base of the OSS-DL.

Each digital library is unique in terms of its content, organization, patrons and services. Hence these criteria would help DL managers make a decision of what tool would be best suitable in their environment. It is therefore not advisable to make across-the-board decision of one tool or the other to all digital libraries.

### 13.6 Future Work

In the present study, more emphasis was given on evaluating selected open source digital library software against the set of defined evaluation criteria. The future of Digital Libraries will be with large amount of data size with varied data types and with many advanced features. It would be therefore worthwhile to examine importing/exporting feature with large volumes of data import/export amongst each software.
Software testing is generally carried out in commercial software environment. There are few open source tools which are available for testing different functions of the software, these tools can be used for testing different functionalities, scalability and load balancing aspects of each OSS-DL. This is an another important and vast area which can be carried out on large severs on these software.

In the present work usability testing of different software have been carried out at very preliminary level but more advanced usability testing can be planned on these software with different usability testing criteria such as 'response time for handling same query', 'how many steps are required to find same information in each software', 'labelling of different terms used in each software', 'visual appearance' etc. Many such studies can be carried out which effectively be helpful to make user interfaces of these software better for user communities.

Many open source digital preservation tools are available under open source license terms and conditions, these tools can be evaluated and a study can be carried out on how these tools can be integrated with the existing OSS-DL which is one of the important area of research.

Creating open access repositories will eventually become a social responsibility. In Europe, UK and USA many academic institutions’ and staff working in those organizations are being encouraged to ‘deposit’ their published papers and other digital documents in an institutional repository. Hence the future will be with more digital repositories with the use of open source software, open standards and open formats. Since different software have different features and different users have different requirements, there is not a universal digital library software architecture that supports everything.
Bibliography


