CHAPTER V
OPEN SOURCE SOFTWARE FOR LIBRARIES

As discussed in the preceding chapter, there are many open source software available for libraries. A wide variety of open source library management software, digital library software and other software are available for the libraries to choose from. This chapter attempts to fulfil the second objective of the present study; thus covers the analysis of open source library management software and digital library software. However, as stated earlier in Chapter III, the researcher has not carried out any primary data collection technique through questionnaire or interview to analyse the various software, but has depended on his personal experience of having worked with some of these software, the manuals and websites of these software and the existing studies on them.

In libraries mainly two software are used, first a library management software known by the generic term as Integrated Library System (ILS) and a software used to create a digital library or an institutional repository. The present chapter would first discuss the open source library management software then move to digital library software. Additionally, this chapter would also discuss some other open source software that can be useful to libraries.

5.1. Open Source Library Management Software

There are many open source library management software in existence, i.e., Avanti, Evergreen¹, PhpMyBibli², PhpMyLibrary³, Koha⁴, NewGenLib⁵, etc. The researcher has

¹http://evergreen-ils.org/
²http://sourceforge.net/projects/phpmybibli
³http://sourceforge.net/projects/phpmylibrary/
⁴http://koha-community.org/
⁵http://www.verussolutions.biz/
discussed them briefly in the earlier chapter. However, out of all these only three software are used by a majority of libraries. These software are Koha, Evergreen and Newgenlib. This current section will discuss only these three software as other software have not made a mark because many of them have either stagnated with a particular version or their development process is very slow or they have unimpressive features. However Koha, Newgenlib and Evergreen could successfully attract the libraries and find enough number of users. A wide number of libraries are using these open source software. Koha and Evergreen are holding about 14 percent market of commercial support ILS services in the world (Breeding, 2013). In addition to this, the number is huge for those who are running open source software on their own without a commercial support. Newgenlib, on the other hand is a developing open source software which originated in India. It is gaining momentum with its excellent features and magnificent support service. Hence it is justifiable to have an analytical and evaluative view on only these three open source ILS which is presented in this section by discussing their technical features and modules.

5.1.1. Koha

Koha, synonymous for the term gift in Maori language, was developed in 1999 to get rid of Y2K problem and on dissatisfactory services from commercial ILS. In Chapter four the investigator has already discussed the historical background of Koha. Koha is written with Perl language and is created on LAMP (Linux, Apache, MySql, PHP) platform. However, a few versions of Koha were also developed for Windows. However, the last version for windows was released in 2009, i.e., 2.2.9 and no further version of Koha was developed for Windows. Koha has won several awards such as 3M Award for Innovation in Libraries in 2000, Interactive NZ Award (The best example of interactive media specifically made for community and not for profit organisations) in 2000, Trophee Du Libre Award in 2003,
Computerworld excellence in IT for non-profit Award in 2004 etc. (Koha, 2013). The current version of Koha is 3.12.4 released on 24th August 2013. The technical specifications and various modules of Koha are discussed below.

5.1.1.1. Technical Features

Besides developed on LAMP platform Koha has several technical features which are as mentioned below in this Section.

**Complete open source**: Koha is a full-fledged open source software. From basic application of Koha to database application, Web server and installation platform all are open source. Koha does not use any commercial software that may levy any indirect costs to a library that adopts Koha. It is distributed under GNU’s General Public Licence while other applications are also distributed under this or other open source licence.

**Library Standard compliant**: Koha is built using maximum library standards and protocols to ensure interoperability between Koha and other software using these standards. Library standards make it easy to migrate to and from Koha easily. Koha Uses MARC tags, ISBD for cataloguing, Z 39.50 for information exchange and several other similar library standards.

**Web based interface**: Koha is not only a Web based application but its interfaces such as administrative interface, OPAC, and self-checkout interfaces are completely web based and opens with the help of a browser. Hence it does not require having a separate client application for running the software. Koha’s web interface is based on XHTML, CSS and JavaScript Web technologies that make Koha platform independent.

**Freedom to use commercial database**: Though Koha is basically tested on MySQL to store data, one may also use commercial database software such as Oracle, or Informix.
No vendor lock in: Koha does not impose any restriction in any form for its use. A library can freely use Koha by itself, with in-house expertise or can outsource external technical support.

Dual database design: Koha uses dual database, i.e., text based and RDBMS to ensure scalable database enough to meet the transaction load of any library, no matter what is the size.

No huge structure required: To have Koha in a library, it does not require to have a huge technical structure. A simple computer is enough to install Koha on it. Koha can work on a standalone computer as well as on a huge intranet and Web structure.

5.1.1.2. Koha Modules

Koha has all essential modules that a complete ILS should have. Koha has acquisition, circulation, cataloguing, OPAC, and serials module. The brief details of the same are as mentioned below in this Section.

Acquisition: Acquisition module of Koha works in two modes. Simple mode and advanced mode. The simple mode provides an interface for adding new holdings to the catalogue. Advanced mode provides options and interface for tracking of acquisition process such as requests, orders, claiming, invoicing, budget control and other processing of the library.

Cataloguing: The cataloguing module of Koha follows worldwide recognized MARC/MARC21 format for creating bibliographic records. Additionally, it can also convert these records to ISBD format. Z39.50 protocol is implemented to retrieve the cataloguing records of other libraries. Cataloguing module of Koha is also compatible with AACR2. Users can create bibliographic record of a document without following acquisition module or
by following acquisition module. Cataloguing module also enables the user to create cataloguing framework for different kinds of material such as monograph, serial, web resources, digital resources etc.

**Circulation:** Circulation module of Koha provides an easy interface for issuing or returning of books. Moreover patron management is also available in this module. This module calculates fine on overdue materials automatically. There is provision to operate with the institutional email server to provide email notifications to the user regarding checking out and in.

**Serials:** Serial module of Koha provides interface for registering periodical subscription, to renew them and to track the arrival of them. It also takes care of late received issues, skipped issues and forwards automatic claims to the supplier. Serials module, in short, is a great hand for the library professionals to manage periodicals with more than ten diverse frequencies. It also forwards the information of library holdings to the OPAC and keeps the patrons abreast about the serial issues available in the library.

**OPAC:** Online Public Access Catalogue of Koha is very interesting. It provides simple Google type search option with an advanced Boolean logic based search. Simple search is for users who are not very familiar with Boolean logic and like to carry a Google like search. It also has ‘did you mean’ feature that lets the user informed about their spelling mistakes or the books available in the database with slight different spellings. Advanced search on the other hand uses Boolean logic to reduce the recall and increase precision. OPAC module also keeps the users informed about the new arrivals in the library. It also enables users to create public and private list of favourite titles. Koha OPAC enables the user to search in various languages. Recently OPAC has integrated with various Web 2.0 features.
Others: In addition to above basic modules, Koha also provides comprehensive report generation facility. Koha can generate a wide range of reports depending upon the requirement of the library. Moreover Koha also helps the library to create barcode labels for books and patrons. It also takes care of Inventory through its inventory management system that helps in stock taking.

Koha, thus, is a complete library management system with all required modules and features. A library can adopt Koha with its own technical human resource or can outsource the maintenance of it with a number of support agencies. The best feature is that the library can change the support agency at any time and it cannot be compelled to take service of a particular agency. Additionally, the libraries that do not wish to take commercial support for Koha may refer a number of manuals available at Koha website, can become a member of the various forums on Koha that extends support to all new Koha users. Moreover, in case of a bug, Koha has created a bug directory that helps to debug any particular problem. Using Koha, a library may get rid of the commercial ILS which often comes loaded with a number of terms and conditions.

5.1.2. Evergreen

Evergreen is yet another prominent open source software which was first released in 2006 and is now adopted by more than 1000 libraries including public libraries, academic libraries, and special libraries all over the world (Evergreen, 2013). It was initiated as a project at Georgia Public Library to take care of more than 200 public libraries of Georgia State. Chapter four of this report has already discussed developmental history of Evergreen in more detail. Evergreen is designed and developed to install at Linux operating system like Koha, however, it is based on client server technology and its client application can be installed on
Windows platform. The current version of Evergreen is 2.4.1 released in July 2013. The technical specification and modules of Evergreen are as below.

5.1.2.1. Technical features

Similar to Koha, Evergreen is also designed using Perl language and uses Apache Web server. However it uses PostgreSQL to store database. The technical features of Evergreen are discussed in this Section.

**True open source:** Evergreen is a true open source software. It is not only designed to install on open source operating platform but all supporting software that it uses are open source only. Evergreen uses open source Web Server Apache, database PostgreSQL, Perl etc. It is licensed under GNU’s GPL licence that offers freedom to users to use, modify and distribute to others.

**Client Server Technology:** Evergreen is built on Client server technology. The server is Linux while the Client application is platform independent and can be installed even on Windows or Mac.

**Library Standards** compliant: Evergreen follows most of the library standards such as MARC, Dublin Core, 739 50 etc. to facilitate easy interoperability between other library tools.

**OpenSRF:** Evergreen also features the Open Scalable Request Framework (OpenSRF) that allows the developers to create applications for evergreen with a minimum knowledge of its structure.

**Applicable in any organisational structure:** Evergreen is compatible to work in any type of organisational structure such as a single library, multiple libraries in the same campus, multiple libraries in various parts of the city or town, or multiple libraries at various cities, states or countries. Evergreen provides complete structure to work in any of these conditions.
**Sufficient for multiple branches:** Evergreen can work for a standalone library or a number of libraries of an organisation. Evergreen under this feature allows each library to set up their own rules and policies within their system. Such as circulation period, number of items issued to the patron etc. This functionality allows each member library to maintain their rules, regulations and individual preferences that makes the workflow of each library very simple and smooth.

**5.1.2.2. Evergreen Modules**

**Acquisition:** Acquisition module in Evergreen has all features and options required for a library. Evergreen provides platform to users to file their request for an item as the library admin can approve or reject a request, create orders for requested book, allot budget for a financial year, receive invoices from vendors and forward it to the technical section for further processing. Moreover, in acquisition module admin can send reminders for items not received from the vendors and can print any document related to acquisition process.

**Cataloguing:** Cataloguing in Evergreen supports MARC format for entering bibliographic details of an item. Admin can modify or create templates using MARC tags. In order to import the cataloguing records from other libraries Z39.50 features is available in this module.

**Circulation:** Through circulation interface admin can issue books to users and can get them checked in back to the library. To specify the period of circulation, admin has to specify the circulation period for various category of users. Moreover, admin can specify the amount of fine for overdue items. Nevertheless, through pre-specified rules admin can select a specific date for any item to any user at the time of checking out of a book. In addition to check in and out, the circulation module provides option to renew the status of an item which is already a
checked out item; option to mark the item as lost or set their status in maintenance in case the book is not available in stacks due to binding or other technical reasons.

**OPAC:** Evergreen OPAC provides a state-of-the-art web interface (Yang & Hofmann, 2010). It provides a simple and Boolean logic based search option to the users. Evergreen provides an excellent navigation from one search result to another with a specific subject or author. Moreover, evergreen provides an optimum ‘did you mean’ feature. When a user’s search hits no result, then Evergreen provides a list of suggestive alternative spelling. It also has the ability to fetch the book cover from Internet on the basis of its ISBN number. If the book is present in Google Books project, then it also hyperlinks to the web page of that particular item so that the users can read the summary of the item.

**Serials:** Serial module is one which was added much later than the other modules in Evergreen. It first appeared in Evergreen in 2009 while full-fledged serial module was added in 2011. Hence, serial module of Evergreen is not as matured as the other open source library management software, however, it has all the basic features such as processing request for a new journal, adding a new journal and a journal issue, registering the claim for a not received issue etc.

**Others:** In addition to the above primary modules, Evergreen has ‘Report Builder’ feature that allows the admin to create various statistical reports using a number of database tables of Evergreen. Moreover like all ILS, Evergreen also has a common administration module to configure the various parameter of the software. To support the users of Evergreen there are a number of online forums run by the community. In addition to that, some commercial vendors are also available for those who wish to adopt a third party commercial support for Evergreen. Some of these vendors are Equinox, ByWater and Lyrasis.
5.1.3. NewGenLib

NewGenLib is another prominent open source software for libraries. It is born and developed in India by Verus Solution with the help of Kesavan Institute of Information and Knowledge Management, Hyderabad. The developmental history of NewGenLib has already been discussed in Chapter IV. NewGenLib is an ILS which is turned open source from a commercial ILS. It received huge publicity in a very short span of time and is being used in a number of libraries not only in India but also overseas. NewGenLib can be installed on Linux as well as on Windows. It is a web based application and based on Client Server structure, however, it does not require any specific client application to be installed specifically. It installs client application automatically from the server. The current version of NewGenLib is 3.0.4 Release 2 released on 25th June 2013.

5.1.3.1. Technical Features

Unlike Koha and Evergreen, NewGenLib has the ability to work on Windows platform in addition to Linux. It uses Java Tomcat as web server, PostgreSQL database to store data and is programmed with the help of Java. NewGenLib has a number of advanced technical features which are rare in any ILS, open source or proprietary. Some of them are discussed here in this Section.

Complete open source: Similar to Koha and Evergreen, NewGenLib (NGL) is also a complete open source software as it uses all open source software to support the functionality of this software. Besides the main application, NGL uses Apache Tomcat as web server, PostgreSQL database and is programmed in Java. It uses Java Web Start to install client application on other computer.
**Platform independent:** This feature makes the NGL unique in comparison to Koha and Evergreen. NGL is platform independent and can be installed on Windows too. It does not compel the users to have a Linux platform to install NGL.

**Library standard** compliant: NGL is complying with international library standards such as MARC, Z39.50, MARC21, and OAI-PMH, Dublin Core. These standards enable interoperability in NGL with other library software or tools.

**RFID integration:** NGL is compatible to RFID (Radio Frequency Identification) integration. If a library wishes to have RFID then they need not change their ILS. NGL is fully compatible to configure RFID based product with it.

**Automated mailing and SMS:** NGL has integrated automated mailing and SMS technology through which a user receives an SMS when s/he borrows or returns the book to the library. In addition to circulation, email and SMS can be sent through various modules such as serials, acquisition etc.

**Multi lingual:** NGL has made it possible to enter data in catalogue in several scripts. Besides English one can enter data in many other languages such as Hindi, Gujarati, Arabic, Urdu, etc.

**Android application:** NGL has developed a mobile application for android so that the users may access NGL through their android mobile (Verus Solutions, n.d.). All basic features of an OPAC are available through this application such as browsing the collection, checking the availability of the item, checking checkout history, list of new arrival books etc.

### 5.1.3.2. NewGenLib Modules

NGL has all modules which are generally found in all ILS and there are some modules which are peculiar to it. The details of the modules of NGL are as follows:
**Acquisition:** acquisition module in NGL provides comprehensive functionality for acquisition. Request made through OPAC are shown here. Additionally the librarian can also add or remove items through this module. The user can create firm order, receive invoices, and assign accession numbers from this module. In addition to the general features of a typical acquisition module this module also manages items received in gift, books bought by the library on approval basis, quotation received by the suppliers etc.

**Cataloguing:** NGL like Koha and Evergreen follows MARC format for bibliographic entry. However, it has eased the work of library professionals by creating a number of MARC based templates for various category of items such as single volume book, multi volume book, book chapter, non-print material, serials, journal article etc. These templates reduce the users' confusion about numerous tags of MARC records. In addition to the ready templates NGL has created several masters which are preconfigured, i.e., type of material, nature of documents etc. Moreover, NGL has provided the functionality to attach a digital file with a MARC record, hence, a library which is not in a huge requirement of a digital library software can do their work with NGL. For a digital document it also follows international standard Dublin Core. Through this module subject and author authority files are stored in the database that increases the vocabulary control in indexing the items.

**Circulation:** circulation of NGL is very simple and communicative. The circulation module works on the basis of the parameters set by the admin such as user privileges and overdue charges etc. Issue and return, and renewal function are very rapid in NGL and it also accepts the barcode standards and RFID integration. Moreover, this module also deals with the items lost or damaged, items sent out for binding, inter library loan, Book Bank, and weeding out of the books.
**OPAC:** Online Public Access Catalogue of NGL is developed with several features of Web 2.0. NGL OPAC provides several search options besides simple and Boolean logic based search. Patron may select the category of items, such from books, serials, book chapters and serial articles etc. Patron by logging onto the OPAC may check books borrowed by them, their privileges and can also request the books for acquisition to the library. In the recently released version a patron may also search the experts in the organisation.

**Serials:** Like other modules, serial module of NGL is also integrated with all basic functions. All functions of Serial desk such as collecting request, creating orders, receipt of invoices and issues, claims for not received issues and binding of serials can be managed with this module.

**NGL Workbench:** NewGenLib work bench provides option to the admin to create and print library card and barcode labels. The library card is fully customizable on the requirement of the library. NGL uses another open source application - iReport for creating library cards using the library admin that can modify the format of library card. Moreover this module also helps the library to import the book record and patron record from an xls or csv file. Book records can also be imported from a MARC file.

**End of the Day Process (EOD):** NewGenLib end of the day process is an application which is bundled in NGL application. The purpose of this application is to relegate the production of time taking outputs, i.e., creating claim letters, emails to users, etc. to a time when the server is comparatively free than the other busier periods of the day.

**Others:** In addition to all modules elaborated above, NGL has an administration module where all basic configurations are done. Through this module various parameters such as holidays, category of patron, list of vendors, budget head, financial year, user privileges, etc. can be fixed. Moreover, NGL has started using a new application that helps to index full text
of digital attachments so that a user can retrieve any documents from any key word of the
document. In addition to the main application, NGL has developed some other applications
which can be integrated with the NGL primary application. These are NGL UC and NGL
Touch. NGL UC (NewGenLib Union Catalogue) helps to integrate the OPAC of several
libraries into one, while NGL Touch is an application that enables the self-check in, check-
out and self-renewal of a library material. NGL also supports VuFind that enables the patron
of a library to search and browse the library resources by replacing the traditional OPAC to
an online OPAC of next generation.

However, presently the users of NewGenLib are limited to many Indian libraries and very
few libraries outside India in comparison of Koha and Evergreen. However, the future of
NGL is very bright looking at its technical support and development process. NGL has
several layers of support to the users of NGL. Firstly, they help the users in real time when
they face a problem or get into a trouble. Using remote desktop access tools, an NGL
programmer may fix any issue in no time. This layer of their support looks more promising
than even technical support of a commercial software. Secondly, like many other software,
they also have a forum where any query may be answered by NGL programmers and the
other users of NGL. In the third option they have created free lancers who may help the
libraries in their area. In addition to all these free support, they may also provide a
commercial support if the library wishes to have it. However, till now, no other commercial
vendor has been observed providing any commercial support for NGL.

5.2. Open Source Digital Library Software

With the explosion in the number of electronic or digital documents, libraries started to store
the electronic copies of theses, dissertations, articles and chapters written by faculty members
and many other forms of scholarly content online for instant and easy access by the users of
the library. A digital library software played an important role in managing, retrieving and
browsing of the electronic documents kept in the server of the library. Many open source software are available for managing a digital library or institutional repository of an institution. However, out of many, three are used by large number of libraries and these are Dspace, Greenstone and Eprints. This section would cover only these three most prominent open source digital library software. However, there are a number of scholarly writings, evaluating the other open source digital library software. For instance, Sunita S. Barve (2008) has done a comprehensive evaluation of open source digital library software in her thesis.

5.2.1. Dspace:

Dspace was developed at MIT, as discussed in the preceding chapter, with the objective to facilitate organisations in managing their digital content. It is a simple and platform independent open source software licensed under BSD Licence (DSpace, n.d.). It is specially designed to support digital preservation for the electronic/digital documents added to the repository in an easy manner (Barve, 2008). The central idea behind the structural model of Dspace is ‘communities’ that is unit and subunit in an organisation. In case of university, ‘communities’ can be schools, departments, laboratories, research centres, etc. of the university (Smith, 2003). Communities contain collections by clustering related content. System architecture of DSpace is as shown in Figure 1.0.

---

6http://www.dspace.org/
7http://www.greenstone.org/
8http://www.eprints.org/
Figure 1.0: DSpace system architecture
(Source: http://www.dspace.org/images/stories/dspace-diagram.pdf)

5.2.1.1. Technical Features

In addition to being an open source software, DSpace uses advanced techniques to make DSpace more advanced in its usability. The technical features of DSpace are as follows:

Complete open source: DSpace is a complete open source software. Besides primary application, it uses all other supporting software such as Java, Apache Ant, Apache Maven, PostgreSQL database and Apache Tomcat; the web server is also an open source. Hence it does not cause any direct or indirect cost on software.

User Interface: DSpace is a complete web based solution. Both interfaces, i.e., user interface and administrative interface, are accessible from any web browser. This helps the
administrator to manage the digital library more effectively. It does not require to access server to manage the collection and changing the preferences and configurations.

**Platform independent:** DSpace is platform independent. It can be installed on Linux as well as any version of Windows. A user need not learn Linux operating system to get DSpace installed, but, it is easily installable on any version of Windows such as Windows XP, Windows Vista, Windows 7, Windows 8, etc.

**Compatible with International Standards:** Dspace is compatible with International standards such as OAI-PMH, Dublin Core, OAICat, etc. Open Archive Initiative- Protocol for Metadata Harvesting (OAI-PMH) is an internationally recognized protocol for metadata harvesting. Dublin core is a set of 15 metadata elements to describe the digital resource (Wikipedia, 2013) and OAICat is OCLC’s open source framework to conform OAI PMH (OCLC, 2011). Moreover, Dspace is also compatible and configured to use any crosswalk plug-in to offer additional metadata formats such as METS.

**CNRI Handling system:** In order to create a tenacious identifier for items uploaded to DSpace, it has used CNRI handling technology. A Handle system is a technological specification for assigning, managing and resolving a persistent identifier for digital objects over the internet (Wikipedia, 2013). Under this specification Handle server runs under a separate process and receives TCP requests from other servers and issues resolution requests to a global server or servers if a Handle entered locally does not correspond to some local content (Handle System, 2012).

**Open URL Support:** Dspace uses Open URL from SFX. Dspace with the help of SFX server may display an Open URL link on every item page using Dublin Core metadata. Moreover, Dspace can respond to the incoming Open URLs.
5.2.1.2. Functional features

Where use of open source technology makes DSpace a complete open source solution for creating digital library; the use of distinctive specification, protocols, standards and programming makes the workflow and function of DSpace very user friendly and attractive. The functional features provided by DSpace are analysed below in this Section.

Authentication and authorisation: Authentication means verification of users to identify their rights in the repository. Users need to log in to use the repository and their rights over the digital library. Administrator can set the limits to various contributors in respect to access of communities and collections. A user who is not authorised or not logged in can access the documents and serve the rights allotted to an anonymous user. Moreover an anonymous user can register him to the repository and can access the rights allotted to an anonymous user; however, the administrator can upgrade the rights of him.

Similarly, the registered contributors (e-people) are managed through authorisation. The admin can control the registered contributors in regards limiting their rights to access the features, collections and communities in the repositories. The administrator can also limit the number of items received from registered contributors and limit the size of items to be deposited into the repository.

Collection and communities: As discussed earlier, the central idea of development of working module of DSpace is ‘communities’, hence at the configuration point the administration has to create the ‘communities’. ‘Communities’ is referred to the collections and people. Administrator creates communities and under communities several collections are created. Similarly e-groups and e-people may also be created. The items submitted by the
contributors are submitted to the related collection of the community. It helps to provide an efficient browse feature at user’s end.

**Flawless workflow:** The workflow of DSpace is based on the need of the institution; however, typically it is a combination of three flawless steps. Each collection in DSpace may be assigned to e-groups or e-persons. If the collection is not assigned to any e-group or e-person then step 1 is skipped and step 2 and 3 are overruled in the absence of step 1. When a depositor submits a document to the repository, it goes to the e-group of the collection where it has been submitted. The group is notified and the group at step one may accept or reject the submitted document. At step 2, the document is informed to another assigned e-group where it can edit metadata provided by the submitter with the item; however the content of the document cannot be edited. Here also e-group may reject the submission of the item. At step three again item is informed to assigned e-group for verification of metadata and submission to the repository. Submission cannot be rejected at this step; however the metadata can be edited. Figure 2.0 shows the workflow of DSpace (2010).

![Figure 2.0: Submission workflow in DSpace](http://www.dspace.org/1_6_2Documentation/ch02.html#N103C6)

A submitter can submit his document following seven simple steps in submission module. The first three steps are related to the description and metadata of the document. At fourth
step contributor has to upload the document to the repository. Fifth step is verification of
submitted information and of document while sixth step deals with the license information.
The submission completes at seventh step, from where it is forwarded to the e-group, to
which document is related for further process.

**Optimized retrieval:** the system in DSpace provides two ways retrieval, i.e., search and
browse of items available in the repository. In search, it facilitates simple and Boolean logic
search. DSpace search retrieves items by searching keywords provided by searcher in the
Dublin Core or any other set of metadata used by the repository. On the other hand, through
browse, a user can browse the communities created under the repositories and subsequently
collection under any community and an item in the collection.

**Supports wide variety of documents:** DSpace supports a huge range of formats in the
repository that helps in creating various types of repositories such as Video, Audio, Picture,
Multimedia, Text etc. Moreover, DSpace also allows the customisation to accommodate the
formats required as per the need of an organisation.

**Import and Export:** DSpace includes the batch tools to import and export the whole data or
a collection of a community. The data is exported in directory form where the metadata is
stored in an XML file.

**Statistics:** Use and user statistics is an attractive feature of DSpace. Through this feature
various reports about the content and user of the system can be generated by the system. The
reports include number of items archived, number of bit streams viewed, number of item
pages viewed, number of collection pages viewed, number of user logins, number of searches
made, most popular search, mostly downloaded item, top country/city views, etc. (DSpace,
2010).
**Others:** In addition to the above features, DSpace has several other small features that are very helpful for users and make the software very valuable. DSpace supports submission of HTML documents to the repository. DSpace also provides support to the Creative Commons licence to be attached to the repository. A submitter may use this license for the item uploaded by him. When a user uses this option, a text document of the creative commons licence is attached with the document submitted by the submitter. DSpace has the RSS feeds for the communities and the collections. When a user subscribes the RSS feed for a particular community or collection, the user gets information through email on submission of a new document to that community of collection. Moreover, DSpace has inbuilt checksum checker to verify if any document in the repository has not become corrupted or has been tampered with.

DSpace is supported by various online forums. These forums have huge number of members from every corner of the world who are using DSpace. Any new user who wishes to adopt DSpace may get help of these forums. In addition to DSpace software, its website also provides complete manual of the DSpace that covers information from installation to configuration and management. However, if an organisation wishes to have commercial support for DSpace, there are wide number of agencies that provide commercial support for DSpace. A list of registered technical support providers is available on DSpace website (DSpace, n.d.).

DSpace is used by numerous libraries and organisations worldwide. It is upgraded at regular intervals. The current version of DSpace is 3.2 that was released on 24th July 2013.
5.2.2. Greenstone

Greenstone is world’s first open source and prominent digital library software. It was developed in New Zealand at University of Waikato with the funding support of UNESCO and Human Info NGO, Belgium (Greenstone, 2012). The historical details of Greenstone are already discussed in Chapter IV. Greenstone is a suite of software that provides a new way to organize information and publish it on the Internet or on CD ROM (Greenstone, n.d.). The aim of Greenstone is to enable users in libraries and other organisations to build their own digital library. Figure 3.0 explains its data model (Trambu, 2012).

Figure 3.0: Data model of Greenstone
5.2.2.1. Technical features

Among all open source digital library software, Greenstone makes itself distinguished due to the technology used by it and a host of features it provides. The technical features of Greenstone are discussed in this Section.

Truly open source software: Like many other open source software, Greenstone also follows a complete open source model. It is programmed using Perl programming language. For web server it uses Apache open source server configured on JRE environment. To store data it uses GDBM (GNU Database Manager) database. All these software are open source software, hence do not attract any hidden or indirect costs.

Supports Windows: While many open source software can be installed only on Linux, Greenstone can be installed not only on Linux, but also on Windows platform. One can install it on 32bit Windows system. It supports Windows XP, Vista and Windows 7.

User Interface: Greenstone works on Client Server technology. Clients can be installed on Linux, Mac or Windows. Installation of Client enables to submit the document from a remote computer. It requires Java to install a client. However the public interface is web based and does not require any specific application. Public interface can be accessed via any web browser.

Single click installation: One of its best technical features is its installation. The new version of Greenstone is bundled in a single software and it does not require long manuals for installation of Greenstone. Additionally, downloading various matching version supporting software is also not required as all supporting software are bundled in one. It is a single click software. Greenstone is just a single exe file, one has to click it once and wait to complete its installation.
**Multilingual**: Greenstone is a multilingual software. It supports Arabic, Chinese, Czech, Dutch, French, Galician, German, Hebrew, Indonesian, Italian, Japanese, Maori, English, Russian, etc. languages. To add any of these languages, one needs to provide a set of language specific text fragments which does not require huge knowledge of Greenstone (Witten, 2003).

**Support to International standards**: Greenstone is compatible with international standards such as OAI-PMH, Z39.50, Dublin Core and other metadata standards. Z39.50 helps to import the metadata of digital documents from the other digital library.

**Large scale database.** Greenstone is designed in such a way that it can contain millions of documents of hundreds of gigabyte size. Even if it is having a huge data uploaded in the server the searching of the server does not get affected.

**Supports plugins**: Greenstone is developed presuming that a number of new formats may be developed in the future and a document may be developed in a format which is not identified by Greenstone. Hence, in Greenstone, plugins can be written to accommodate new types of data.

**5.2.2.2. Functional Features**

Using above mentioned technologies and supporting facilities, Greenstone provides functional facilities that distinguish it from other digital library software. The functional features of Greenstone are analysed here.

**Multimedia**: Greenstone supports a wide variety of documents. The Greenstone collection can contain textual documents, pictures, audio files, video files and a combination of all of these. Hence Greenstone can be very useful software to a variety of institutions, for example,
video library, multimedia library, academic library, media library, etc. Greenstone also enables to write plugin to identify other types of documents which are not facilitated by default.

**Generates CD-ROM databases:** This feature and function of Greenstone separates it from other digital library software. This feature makes it a complete digital library software as one can also create CD-ROM based database. If a library wishes to create subject based collection on CD-ROM, Greenstone provides this facility. The databases created by Greenstone can be installed on any local computer and documents on it could be retrieved from here.

**Flexible searching and browsing:** Greenstone provides search as well as browse option. In search it provides simple and advanced search. It also provides option to search a field or full text of document. In browsing one can browse the collection through author, subject, title, uploaded dates, etc. Browsing options can be selected at the time of building collection.

**Real time operation:** New collection in Greenstone can be installed without bringing the system down. Users who are online may also notice newly installed collection instantly.

**Supports various metadata formats:** Greenstone supports a number of metadata formats such as Dublin Core, RFC1807, NZLS (New Zealand Govt. Location Service), AGLS (Australian Government Locator Service) and many more. These metadata formats can be used by the organisations as per their needs. Moreover, an organisation also can create their own metadata format by choosing and creating metadata elements. It facilitates the organisation to manage various online databases such as video, audio, patent, newspaper clippings and multimedia effectively.

**What you see-what you get (WYSWYG):** Greenstone is an open source digital library software, available from the New Zealand Digital Library (ngdl.org) under the terms and
condition of GNU General Public Licence. It includes everything discussed in this section such as web access, capacity to generate CD-ROM based databases, etc. (Witten, Bainbridge, & Boddie, 2001). It also provides an auto install feature for easy installation on Windows and Linux.

**Others:** In addition to the above mentioned functional features, Greenstone provides various small features such as it can be installed offline also. Being multilingual it is supported by International communities in more than one language. It organises the document in hierarchical form that one may browse the collection in hierarchical way. Greenstone also facilitates distributed collections. Distributed collection means various collections installed on various computers, which may be presented to the users in the same way and through a single web page as a part of a single digital library.

Greenstone is supported by large community to provide technical assistance to its users worldwide. It has forums, Wikis, mailing lists, FAQs, user manuals, to support the users. Additionally, there are also a number of online training programmes organised by its developers and the community to inform users the features of Greenstone. Moreover, a number of conferences are also held to discuss Greenstone and workshops are conducted to provide practical hands on training to it. In addition to all, commercial support for Greenstone is also provided by a number of vendors worldwide. Greenstone is used by a huge number of organisations worldwide. The current version of Greenstone is 2.86.

### 5.2.3. EPrints

EPrints is another digital library software on the lines of DSpace and Greenstone. It is yet another highly used open source software for creating digital library after the above two. Eprints took its way at Southampton University in 2000. The historical background of EPrints
is discussed in Chapter IV. Originally the development of EPrints was supported by Cogprints, and presently it is supported by Joint Information Systems Committee (JISC) as a part of Open Citation Project. The technical features of EPrints are as below:

5.2.3.1. Technical features

Like many other open source software, EPrints is also an open source application and is distributed through the web without any cost and with liberty to modify, redistribute and adapt as per the need. However there are several features that differentiate EPrints from other open source digital library software. The technical features of EPrints are as follows.

True Open source based on LAMP technology: EPrints is developed on the LAMP architecture, however it is written on Perl instead of PHP (EPrints, 2013). All software used for EPrints are fully open source software and thus they do not attract any hidden or indirect costs to adopt it. It is licensed under GNU’s General Public Licence which is a prominent licence for open source software.

Platform Independent: With a release in 2010 for Windows, EPrints has become a platform independent software. Now it can be installed on Linux, Mac or Windows. Linux is not a must operating system for using EPrints. A user of Windows operating system can also adopt EPrints.

User Interface: Unlike Greenstone, EPrints provides web based interface not only for user but also for the administrator that makes it easy for the administrator to manage the software from any computer on LAN or on WAN.

Supports Plugins: Being written in Perl, it is very easy to write Plug-ins for EPrints. It also supports third party plug-ins to get compatible with unknown formats.
Complies with International Standards: EPrints complies with international standards such as OAI-PMH and various Metadata standards. Use of these International standards makes EPrints compatible with other software and increase the interoperability of the software.

Multilingual: EPrints has used Unicode thoroughly; hence it may allow use of any language consistently.

Export and import: EPrints facilitates bulk import and export of record in several formats such as ASCII, BibTex, XML, etc. that helps to create large collection easily.

5.2.3.2. Functional Features
EPrints has a different working structure than that of DSpace and Greenstone. The functional overview of EPrints is as follows:

Three user role: EPrints has its work flow around three major roles, i.e., author, editor and administrator (Beazley, 2010). Author has to submit his/her documents or papers, editor’s role is to review the submission before publication through the repository and administrator controls all back-end functions of the software that includes technical and functional structure.

Supports wide variety of media: EPrints supports a variety of formats such as PDF, JPEG, TIFF, HTML, etc. Those formats which are not supported by default can be recognized by using plug-ins.

Extended search option: EPrints provides extended search option that includes simple search, field search, and full text search. A search in EPrints scans through each metadata type entered in the database. In addition to the search option, EPrints also provides browsing
facility through which one can browse the collection on the basis of author, title, subject or date of submission.

**Customisable metadata:** EPrints supports a number of metadata schemas such as Dublin Core, METS, etc. In addition to this it also supports the customisation of the metadata that facilitates the creation of a metadata by defining the elements as per the requirement of the organisation.

**SHEPRA/RoMEO Integration:** EPrints provides integration with SHEPRA/RoMEO for quickly checking the publisher’s policies and author’s rights over the content submitted by them (Beazley, 2010).

**Others:** In addition to the above functional features, EPrints has many small features that attracted the large community to use it. EPrints provides RSS feeds for entire collection based on the specific criteria decided by the administrator such as author, subject, etc. Having a web interface has eased the end user submission to the repository. It supports Library of Congress Subject Heading for indexing and browsing of the collection.

At present EPrints has been reported in more than 270 organisations including British Library, Harvard College, Australia National University, etc. Like other open source software discussed above, EPrints also has a similar support. One may use the ready documents on EPrints website, EPrints forums available on the web or can use the option of commercial support. A very active community is working behind the EPrints for its consistent development. The current version of EPrints is version 3.3.12 released on 24th July 2013 (EPrints, 2013a).
5.3. **Other open source software for libraries**

In addition to open source library management software and open source digital library software, there are a number of software available that can be used by a library. These software include Content management software, Journal management software, etc. Engard (2010) has listed a wide range of open source software in her book on practical open source software. She has also listed the experience of some libraries using those open source software. Some categories of open source software not discussed in the earlier sections that may be very useful in libraries are discussed in this Section.

5.3.1. **Open Source Web OPAC (Online Public Access Catalogue)**

When we say Web OPAC, an OPAC bundled in library management software reflects in our mind. However, here the concern is not about the default Web OPACs of library management software; open source web OPAC are the applications, integrated with Web 2.0 features, designed especially for converting the OPAC into a next generation Web OPAC without modifying the existing or default Web OPAC application of the library management software. These applications provide an additional interface for the users to retrieve the library resources with the help of a third-party application known as catalogue overlay system or Catalogue 2.0. Some of the very prominent such OPACs are Acquabrowser, Scriblio, Vufind, SOPAC, Backlight, LibraryFind, Extensible catalogue, ObiblioOPAC for Joomla, OPACIAL, Drupal etc. An extensive comparison of many of these is done by Yang and Wagner (2010). These applications, however, are designed by using diverse structure and means, but the objective is to provide a next generation catalogue. Few of these are solely cataloguing applications while others provides catalogue integration as a feature of their main application, however the main application is not a library integrated system. Some of these applications are used frequently by the libraries such as VuFind, LibraryFind, and Scriblio.
VuFind is a solely web catalogue application integrated with a number of Web 2.0 features and was developed in Villanova University Library in 2007; however, it could detect only four users by July 2009. Using VuFind with any ILS needs bit of programming if it is not already supported by the ILS. The information on technical requirement for implementing VuFind and information about the various features is available on project website of VuFind\(^9\).

VuFind, like other open source software, is released under GNU General Public Licence and available for download from its website. NewGenLib is programmed to adopt VuFind that makes integration of VuFind with NewGenLib very simple. Koha can also be integrated with it with little modification (Adara, 2011). Using VuFind will make the catalogue of library more advanced and dynamic.

LibraryFind\(^10\) is another open source application to make the library catalogue more technical and attractive. It was developed by Oregon State University Library, United States in 2007. LibrayFind also provides various Web 2.0 features for information retrieval. Project site of LibraryFind provides all required information for the installation of this application on any computer. However, there is lack of documents providing information on integration of LibraryFind with Koha, NewGenLib or any other library ILS.

Scriblio\(^11\), earlier known as WPopac, is the oldest open source OPAC with Web 2.0 application developed in 2006 by Host Plymouth State University, United States. Scriblio is fully based on Wordpress, the online content management application. Scriblio not only provides the library an opportunity to create a free website but also to integrate its OPAC with this website. The detailed technical notes on installation and working of Scriblio are available on its project Website.

---

\(^9\) [http://www.vufind.org](http://www.vufind.org)
\(^10\) [http://www.libraryfind.org](http://www.libraryfind.org)
\(^11\) [http://scriblio.net](http://scriblio.net)
ObiblioOPAC for Joomla and Drupal are OPAC modules that can be integrated with Joomla or Drupal to configure a web OPAC on a website designed on Joomla or Drupal. These applications are rather simple than other applications as they are created in the form of a module. However, precondition for these applications is having Joomla or Drupal based website.

Although above applications are just an add-in for the ILS, they require tedious procedures to install and configure them with the ILS. This tedious procedures is one of the reasons for not attracting a huge number of users worldwide.

5.3.2. Open Source Content Management System

A content management system (CMS) is a software providing ability to create a website with dynamic features of Web 2.0 such as blog, RSS feed, discussion forum, surveys, wikies, podcasts, etc. Content management system is of two types: web based and an offline application. The Web based content management system lets the users not only to create website but also hosts them at the real time while an offline application enables users to create a website offline which the user can upload online by choosing host and registering domain name of his/her choice. Both types of options are available through proprietary and open source. However, free web hosting and domain name is not guaranteed in online content management system. Hence, it is always favourable to create a website using offline applications where users have the freedom to select the web hosting and domain name vendors. In offline CMS, a huge number of open source applications are available, however, very few are used by majority of sites. Wordpress\textsuperscript{12}, Joomla\textsuperscript{13} and Drupal\textsuperscript{14} are mostly used and highly referred CMS applications. Out of these three applications Wordpress covers

\begin{itemize}
\item \textsuperscript{12} http://www.wordpress.org
\item \textsuperscript{13} http://www.joomla.org
\item \textsuperscript{14} http://www.drupal.org
\end{itemize}
maximum market than Joomla and Drupal due to its simplicity and ease. Joomla is on second place while Drupal is on third as it is difficult than other two, although it is very high-tech than others.

Joomla, Drupal and Wordpress, all three being open source, are developed on PHP and MySQL. However, they dynamically differ in terms of their capabilities and features.

Wordpress is the simplest CMS of the three listed above. It was released in 2003 under GNU’s GPL license making it complete open source software. One can design a blog or a website using Wordpress easily. However, those who wish to create more dynamic website may prefer Joomla or Drupal. The project page of Wordpress has enough number of document to assist in its installation and using its modules. Wordpress enables users to handle it from multiple points hence it is a multi author CMS tool. It is a user friendly tool having a plugin library to add various features in the website. A beginner may easily opt for it for creation of a library blog or a simple library website. There is also a service, i.e., Wordpress.com, where users not only can create websites but also host it free and get a free domain name. The drawback of Wordpress is its limitations. It has limited designs, limited plug-ins and limited content management capabilities that make Wordpress a good blog creation tool but very basic CMS for the newbies. However, in terms of use, Wordpress is at the top of all three CMS. Some featured websites created with Wordpress are metro.co.uk (Metro, Newspaper, UK), tv.msnbc.com, blogs.loc.gov/loc/ (Library of Congress, blog), library.plymouth.edu (Plymouth University, Library, lis.uncg.edu (University of North Caroline, Library), etc. A comprehensive list of many other sites using Wordpress is given on the project page of Wordpress.

Joomla is another open source CMS and winner of several awards for its wonderful features. It was released in 2005 and is licensed under GNU’s GPL, the open source license. The great
advantage of using Joomla is its strong developer community\textsuperscript{15} that is always ready to support the users of Joomla. Like Wordpress it is a user-friendly CMS tool that requires no technical or programing knowledge to handle it. However, having a technical knowledge is always a benefit in creation of any website or blog. Joomla is a more powerful tool than Wordpress, having many extension tools that include plug-ins, components, templates, modules and languages. Joomla is also supported with a large number of manuals for its installation and modular handling. The weakness of the CMS is that one cannot start working straight forward with Joomla, one needs to learn crunch of it before working on it. However, it helps one to develop a site which has more structural stability and content than Wordpress. Some featured websites created with Joomla aregsas.harvard.edu (Harvard University), statelibrary.sc.gov (South Carolina State Library), www.librarynext.com (LibraryNext, India), etc.ou.okstate.edu (Oklahoma State University) and many more can be explored on Joomla project website.

Drupal is another prominent open source CMS that came into existence in 2001 with its distribution under GNU’s GPL license. Drupal is exceptionally powerful and developer friendly CMS. It is very flexible to develop powerful blogs or a dominant website. It has more features and possibilities than Joomla or Wordpress. Project page of Drupal provides a number of manuals to help in installation and working on Drupal. However, it requires more expertise in using a CMS than Joomla and Wordpress. Drupal with a large number of modules and themes makes itself very useful for making websites with great features. The current version of Drupal is 7.23 released on 7\textsuperscript{th} August 2013. Some featured websites that developed using Drupal are thewhitehouse.gov (White House, USA), data.gov.uk (Govt. of UK), science.mit.edu (MIT School of Science), library.unt.edu (University of North Texas Libraries), and college.du.ac.in (Acharya NarendraDev College, India), etc.

\textsuperscript{15}http://community.joomla.org/
Although there are a number of open source CMS available, but the three discussed above are the most important and prominent of all. Using any one of these three may be useful in all respects, i.e., security, documentation and support.

5.3.3. Open Source Journal Management and Conference Management Software

Public Knowledge Project\textsuperscript{16} is the developer of Open Journal System (OJS), Open Conference System (OCS) and few other software. OJS is an open source software developed with the objective to provide a solution to host and manage all aspects of publishing a journals online such as call for paper, receiving manuscripts, forward them to editors, resending to author for correction, peer review, publication of an issue, and retrieving the current or older issues and journals. OJS made each of these works very easy and convenient. Any organization publishing a periodical can easily manage all operation related to Journal with OJS. It is an open source software released under GNU’s GPL license.

Open Conference System on the other hand, is another important software that is used widely. A conference, symposium, workshop, etc. can be managed and handled with the help of OCS. OCS is compatible to handle every operation of a conference management including publication of conference paper. Like OJS, OCS is also released under GNU’s GPL licence that makes OCS an open source software. In addition to these two software, Public Knowledge Project has developed Open Monograph Press (OMP) and Open Harvester Software (OHS) designed on the basis of OJS and OCS.

All these software are user friendly and can be installed on any platform. Installation of these software is easy. The manuals of working on software are available on its project website.

\textsuperscript{16} http://pkp.sfu.ca
5.4. In Summation

The open source software has given the libraries a chance to provide the users the services comparable to the libraries with huge economic and technological resources. Open source software could be adopted at a very low cost without any vendor lock in and with the source code. Then one can increase the innovation and adaptation of the software by modifying the source code of the software. Open source software are imperative for any type of library, i.e., academic, public or research. The philosophy of open source software revolves around two principles, viz., sharing and collaboration. These principles are very much acceptable and adaptable in libraries too. Hence, the open source software and libraries are natural fit, as both promote learning and understanding by dissemination of information (Poynder, 2001). Open source software come to libraries as exciting opportunities to fulfil the expectations, roles and process which are critical to the library’s community of patrons. These open source software are available in many categories as discussed in this chapter and hundreds of other categories which are important to an individual or organisation at any level. OSS in libraries are available for every function that can embrace the quality of service provided by the libraries. A library having a single computer with internet connection can easily adopt and implement these open source software. Open source software does not require high license fee or high configuration fee for the software implementation in the library. Moreover, open source software also has the possibility to get the commercial services of the vendors, if the library is willing to have it.

Open source software with these opportunities also pose some challenges forward. If an organisation is not having techno savvy staff to implement these software then it may attract a cost by hiring a vendor to implement it. Another challenge is that generally open source does not offer the same kind of support as a proprietary software does. At the time of any error or bug libraries with proprietary software may easily call the vendor and get the
problem solved, but the case is not the same in case of open source software. Here the person dealing with the software has to solve the problem himself or herself. S/he has to rely fully on the community forum of the particular open source software which may take time to solve the bug or error. Open source software can be a daunting game for those who are not having programming knowledge and confidence or are not techno savvy.

However, these challenges are no reason to not choose open source software, but just are the challenges that need to be taken care of. These challenges can be faced even when a library has commercial software. A library must select the software keeping these challenges in mind. There are a number of excellent open source software as discussed in this chapter which are available and are endowed with tremendous features but a library need to keep in mind the resources available with it while choosing the software. If one chooses an open source software keeping in mind these challenges and the available resources (technical expertise) then there can be no better alternative than open source software. The next chapter deals with the optimum software model that a library can adopt.

References


