Enabling ILAP as Digital Library Software: A case study with KOHA

Shri Ram Pandey **Pankaj Kumar Singh**

Abstract

Open Source movement has tremendously influenced Information industry in general and digital libraries in particular. Many of the open source digital library software available on the World Wide Web are developed based on highly sophisticated programming language, which needs in depth knowledge and guidance for installing, implementing, customizing and integrating with other open source software. They generally do not cater to the needs of beginners who want to develop software in open source environment. Present paper is an attempt to enable Integrated Library Automation Software (ILAP) Koha, as a digital library especially koha version 3 in the view that koha can support digital library functionality.

Keywords: Digital Library, KOHA, Digital Library Software

1. Introduction

Open Source has been gaining much attention in the last few years. Many corporations, large and small, have taken an interest in this growing software market that shows some strong differences with traditional software. Open Source is no longer seen as an insignificant niche market but as a serious development in the software market. Digital libraries (DL) have become a major part of the mainstream library landscape, as open source software (OSS) has become a worldwide phenomenon. According to the Digital Library Federation (DLF) digital libraries are:

Organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities (Digital Library Federation, 1998).

The Open Source Digital Library software (OSDLS) offers some great opportunities to develop a digital library. DL are the repository of information, but one can search and browse from any location, any time. DL save time and money by providing around-the-clock solutions to information needs, including emerging and existing technology and business knowledge. DL is now widely recognized as an important component of a global information infrastructure for the new century. Researchers all over the world are concerned about using innovative information technologies for managing and manipulating digitized content. In this changing scenario library and information professionals have to learn more and more about computer systems, file formats and web servers and basic of the programming language to sustain in the current world of open source. Over time, commodity software can hide away certain levels of complexity as a basic infrastructure is developed. But innovative computer applications will always require a degree of explicit support and tinkering. For example, the Greenstone digital library software development team has, over several years, developed a variety of tools to abstract away from technical details to simplify the tasks of creating and maintaining

Shri Ram Pandey and Pankaj Kumar Singh

digital collections. Despite this work the experiences of using digital library software (DLS) can still be disconcerting for some students (Nichols et al, 2006). This state of affairs is more a reflection on the state-of-the-art in content management systems than a criticism of any particular software applications. From the perspective of software developers the power of DLS derives from the flexibility of computer programming languages.

Current developing trends in DL are promoting open access software to online information delivery and disseminated via the Internet and Web. While the field of software is promoting open source software, developing open source digital library also becoming a challenging issue, as it is one of the primary requirements for the further evolution of DL in the twenty-first century. The basic concept of DL is that libraries accessible from anywhere, anytime. The dream of DL has become a reality as a result of rapid advances in technologies.

The present study is a approach to develop DLS using LAMP (Linux, Apache, MySQL and PHP) architecture and integrate it Koha(A Integrated Library Automation Package.)

2. Need of the Study

Since late 1990's, many studies have been made to describe and develop DLS. However, very few attempts have been made by library and information science (LIS) professionals to enabled a library automation package as digital library. To shed some light on preliminary efforts to develop digital library software and integrate it to the library automation software, this study has been carried out. Present work is intended to draw a clear roadmap of enabling a library automation software as digital library to all those information professionals who have a strong interest in developing basic digital library software and integrating with Koha using LAMP architecture.

Such an effort to develop digital library software based on LAMP architecture makes it possible to easily integrate with any software using LAMP architecture and requires a few basic modules of DL to be incorporated.

3. Koha: The First Open Source ILS

Koha is the first open source Integrated Library System (ILS) with true enterprise-class ILS functionalities including circulation, cataloging, acquisitions, serials, reserves, user management, branch relationships, etc. It is built using library standards and protocols to ensure interoperability ith other systems and technologies and provide a platform-independent solution. Developed initially in New Zealand by Katipo Communications Ltd., Koha was first deployed in Jan 2000 for Horowhenua Library Trust and has spread across the world since then. Koha is distributed under the General Public License (GPL) and is currently maintained by a team of software providers and library technologists around the globe.

Koha is the first open-source Integrated Library System (ILS). In use worldwide, its development is steered by a growing community of libraries collaborating to achieve their technology goals. Koha's impressive feature set continues to evolve and expand to meet the needs of its user base.

Koha is available for free download from the Koha Web site or from one of the companies that supports the open source software. These companies, including LibLime in North America, do not charge for the software, but do charge for consulting, programming, training, technical support, and the hosting services they provide.

3.1 Why Koha?

3.1.1 Full-featured ILS. In use worldwide in libraries of all sizes, Koha is a true enterprise-class ILS with comprehensive functionality including basic or advanced options. Koha includes modules for circulation, cataloging, acquisitions, serials, reserves, patron management, branch relationships, and more.

3.1.2 Dual Database Design. Koha uses a dual database design that utilizes the strengths of the two major industry-standard database types (text-based and RDBMS). This design feature ensures that Koha is scalable enough to meet the transaction load of any library, no matter what the size.

3.1.3 Library Standards Compliant. Koha is built using library standards and protocols that ensure interoperability between Koha and other systems and technologies, while supporting existing workflows and tools.

3.1.4 Web-based Interfaces. Koha's OPAC, circ, management and self-checkout interfaces are all based on standards-compliant World Wide Web technologies—XHTML, CSS and Javascript—making Koha a truly platform-independent solution.

3.1.5 Free/Open Source. Koha is distributed under the open-source General Public License (GPL).

2.1.6 No Vendor Lock-in. It is an important part of the open-source promise that there is no vendor lock-in: libraries are free to install and use Koha themselves if the have the in-house expertise or to purchase support or development services from the best available source.

3.2 Why Koha 3.0

Koha 3 has much unique quality as follows:

(1) OPAC

i. Library's own holdings of print and electronic resources;

ii. Federated searching of the catalogues of partner libraries

iii. Federated searching of open access and subscription e-resources, including

iv. Simultaneous management of members' log-in to their account and

v. E-subscriptions, and Athens authentication;

vi. Clustering of search results; and

vii. Readers can save searches, set up alerts, create RSS feeds etc. to their profile when logged in.

(2). Cataloguing: compatibility

i. MARC21;

ii. Z39.50, XML; and

iii. Ability to link to locally and externally hosted online documents in a range of formats.

(3). Circulation

i. Reader data to be bulk-imported from stakeholder organizations;

ii. Members' data to be exported for card printing;

iii. RFID support;

iv. Support for self-issue machines (Standard Interface Protocol (SIP2) or Niso

v. Circulation Interchange Protocol

vi. Administration of fines and courtesy/overdue notices.

(4). Acquisitions and Budget Management

i. Communication with suppliers' databases for books and journals (Electronic Data Interchange (EDI));

ii. Management of binding;

iii. Support for multiple budgets;

iv. Support for different ownership of collections; and

v. Management of library income from fines and sales.

(5). Consortium Functionality

i. LMS installation to support multiple independent libraries

ii. Hardware and operating system-independent browser-based solution preferred.

(6). Hosting

i. Both local and remote hosting options can be considered.

(7). Impact on staff

i. System to be user-friendly and require minimum training and expertise.

3.3 Koha as a Digital Library

Recent advances in information technology have provided new ways of dealing with information in academic libraries. The explosion of the web, the internet and information technology in general has created challenges and has provided huge opportunities for those working in library and information professions in higher education. Open source Automation software could also now equipped with the digital library functionality which provides researchers and students the ability to read full text and research outputs.

Koha is a integrated library automation software, which includes almost all modules which is requires for a Library but it does not have the digital library functionality.

3.4 Approaches to Use Koha as Digital Library

We can adapt two approaches to develop koha as Digital Library. The first approach is to integrate it with some existing digital library software like Greenston, Dspace etc.. Second which is the scope of this paper is to develop a separate model based on LAMP (Linux, Apache , Mysql and PHP) which can support full text and integrate with the Koha.

3.5 Basic Idea and Requirements

Considering this approach in mind three main technologies are required: a web server, a database and a mechanism for these two to interact and incorpo- rate database content in the Index (html pages). We chose to run the Apache web server (see http://www.apache.org), the most popular web server on the Internet2 on a UNIX server. For a relational database, we chose MySQL (http://www.mysql.com) which is stable, has a very fast query response time on simple tables, is reliable and ANSI SQL92 compliant.3 To glue these two together, we chose the increasingly popular server-side HTML embedded scripting language PHP (see http://www.php.net). PHP was initially developed as a personal project, hence the original name Personal Home Page. It is now a major open source server-side HTML embedded scripting language and is in use at over two million websites. PHP is available as an Apache module, and is therefore very efÞ- cient, and works well with MySQL (as well as a variety of other databases). PHP is very easy to integrate with HTML. Blocks of PHP code can simply be inserted in a standard HTML document as required. One of the most useful features of PHP is the way it handles HTML forms. Any form element automatically results in a variable with the same name being created on the target page. All the elements in the form are therefore available as variables and ready to incorporate in your HTML output or include in a database query. PHP can perform mathematical calculations, manipulate strings of data and (as in this case) query databases. It can work with many predeÞned functions, including XML support, e-mail capabilities (including IMAP functions) and automatic session handling.

3.6 Stages to Development and Design Digital Library Software

3.6. 1 Requirements Specifications

One of the first design issues in the creation of digital library software is to prepare a list of high-level requirements. This list includes what information the library contains, how that information is generated, what audience the information is intended for, and how the data will be accessed. All of these issues must be considered in the development of DLS. A clear plan must be developed before one starts the detailed design and development of the library software.

3.6.2 Storage - Database or File Structure

The next big decision to be made is how to store the files that comprise the digital library. There are a number of options available. Two possible methods are using a database and creating a special directory & file structure. There are many databases to choose from. Examples include Oracle, MSQL, and Microsoft Access. If you use a Windows operating system, access would probably be the database of choice, as it would be the most compatible with NT. Oracle is a good choice for a database on a UNIX machine but as objective is to develop the software on LAMP platform, MySQL is selected as back end. It is very reliable and fairly flexible. An E-resource database is created using MySQL.

A. The E-resources database

The E-resource database contains metadata for various resources. The database is constructed in response to a number of specific requirements. The first requirement is that the various resources should be consistently described using a number of metadata elements. These are resource ID, title, author, publisher, year, item type, department, file format. It should be possible to search for resources by title or title keyword. In addition, it should also be possible to browse a list of resources, by title, author and year.

B. Data analysis and Database Creation

In a relational database multiple tables of data relate to each other through special key fields .The main advantage of a relational, as opposed to a table, database is that data duplication, and therefore potential inconsistencies, are eliminated. In response to the requirements, following tables were designed within the E- resources database to contain the metadata. These are outlined below:

- resource - resource details (title, description etc);
- subject - subjects and their introductions;
- format - common formats and access methods
- resource format – tables which formats/access methods apply to each resource.
- year-publication year of that resource
- department-Belongs to particular department
- url for upload the resource in various file formats like PDF, Audio/Video

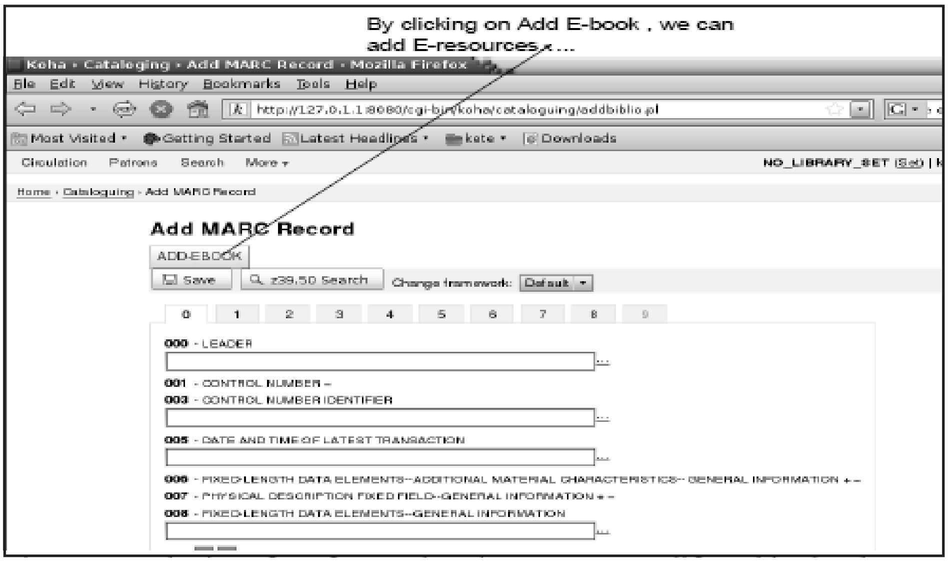
C. Database Creation Command

```
CREATE TABLE `put` (
`id` VARCHAR( 50 ) NOT NULL ,
`title` VARCHAR( 100 ) NOT NULL ,
`afn` VARCHAR( 50 ) NOT NULL ,
`aln` VARCHAR( 50 ) NOT NULL ,
`publisher` VARCHAR( 50 ) NOT NULL ,
`year` VARCHAR( 4 ) NOT NULL ,
`item_type` VARCHAR( 50 ) NOT NULL ,
`department` VARCHAR( 50 ) NOT NULL ,
`item_format` VARCHAR( 20 ) NOT NULL ,
`name` VARCHAR( 40 ) NOT NULL ,
`TYPE` VARCHAR( 30 ) NOT NULL ,
`size` INT( 11 ) NOT NULL ,
`content` MEDIUMBLOB NOT NULL ,
PRIMARY KEY ( `id` )
);
```

4. Stages of Implementation

Many PHP scripts are written for supporting different functionality as given below

A. Administrator Login
B. Cataloguing
C. Searching
D. Browsing by year, title, author etc.



By clicking on Add E-book , we can add E-resources ...

Koha › Cataloging › Add MARC Record - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://127.0.1.1:8080/cgi-bin/koha/cataloguing/addbiblio.pl
Most Visited
Getting Started
Latest Headlines
kete
Downloads
Circulation
Patrons
Search
More
NO_LIBRARY_SET (Set)
Home › Cataloguing › Add MARC Record
ADD-EBOOK
Save
z39.50 Search
0
1
2
3
4
5
6
7
8
9

Add MARC Record

Change framework: Default

000 - LEADER
001 - CONTROL NUMBER –
003 - CONTROL NUMBER IDENTIFIER
005 - DATE AND TIME OF LATEST TRANSACTION
006 - FIXED-LENGTH DATA ELEMENTS--ADDITIONAL MATERIAL CHARACTERISTICS-- GENERAL INFORMATION + –
007 - PHYSICAL DESCRIPTION FIXED FIELD--GENERAL INFORMATION + –
008 - FIXED-LENGTH DATA ELEMENTS--GENERAL INFORMATION

Figure 1: Koha interface for cataloguing we can modify Add-e book

The first thing you we need to do is to modify the scripts of koha. Koha as such have not any functionality to support attachments and to add full text. So in order to make koha to support attachment, it is necessary to develop modules which will support attachments as well as various browsing and searching facility.

This can be done by following ways—

1. Modifying the template addbiblio.tmpl file in koha, so that user can add e-resources during cataloguing.

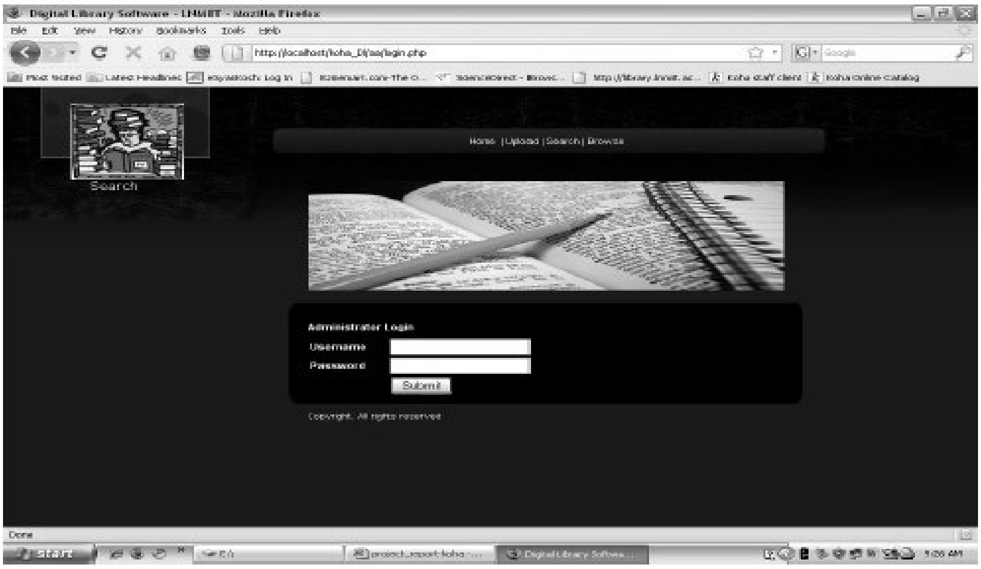
2. Defining link to that separate modules.

After clicking the link E-BOOK (figure-1), it automatically redirected to separate module koha digital library where user can add/catalogue resources, search and browse the documents.

This separate module works basically in three steps.

4.1 Administrative Login

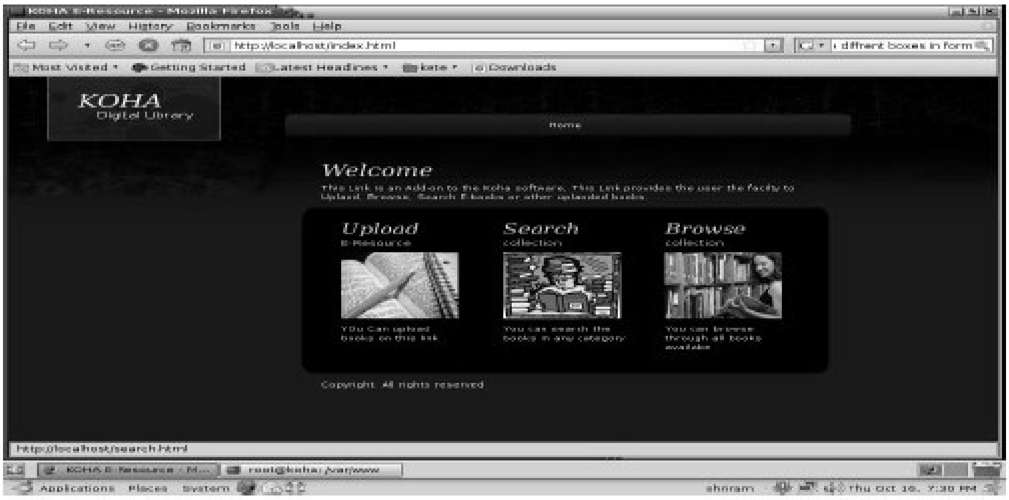
The first interface of this, module is for administrative logging. Where administrator has only privileges to add and catalogue e-resources.


Digital Library Software - LNMIIT - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost/koha_Dl/ss/login.php
Home |Upload |Search| Browse
Search
Done
start

Administrator Login
Username
Password
Submit

Copyright. All rights reserved

This interface also incorporates other functionalities like searching and browsing. For searching and browsing the content there is no need of authentication. After login through as a administrator, user will find many other features like cataloguing, modify and log out options. The idea behind designing this first page is to set preliminary privileges to access and catalogue the e-resources. Figure 2 and 3 shows a screenshot to support the admin login and other features.


KOHA E-Resource - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost/index.html
Most Visited Getting Started Latest Headlines kete Downloads
KOHA
Digital Library
Home
Upload
E-Resource
Search
collection
Browse
collection
http://localhost/search.html
Applications Places System
shriram Thu Oct 16, 7:30 PM

Welcome

This Link is an Add-on to the Koha software. This Link provides the user the facity to Upload, Browse, Search E-books or other uploaded books.

YOu Can upload books on this link

You can search the books in any category

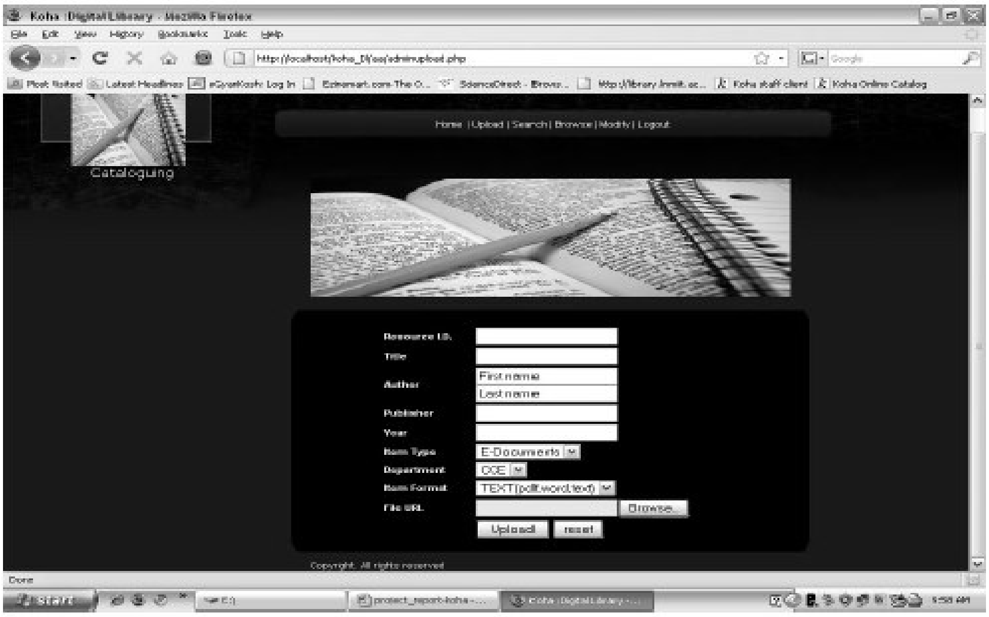
You can browse through all books availabe

Copyright. All rights reserved

Figure 2& 3: The main interface for DL module where user can search/browse etc

4.2 Add/Catalogue

This interface facilitate user to catalogue the e-resources based on standard metadata. It provides facility to add the e-resources and to give some basic metadata like title, Author, Publishes, Department etc. We can also add the full text of the files in various formats like PDF, .AVI, .EXE etc. It supports all the file formats used today. Figure: 4 shows the metadata fields of this function.


Koha :Digital Library - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost/koha_DL/asa/adminupload.php
Cataloguing
Home | Upload | Search | Browse | Modify | Logout
Browse...
Upload
reset
Done
start

Resource I.D.

Title

Author: First name / Last name

Publisher

Year

Item Type: E-Documents

Department: OCE

Item Format: TEXT(pdf,word,text)

File URL.

Copyright. All rights reserved

Figure 4: The main interface for DL module where user can search/browse etc

4.3. Searching: This feature facilitates the user to search the documents by metadata fields such as Author, title, resource-id/acc-no and department. This interface also provides the facility to further access the resources alphabetically. There is predefined structure to display the results. Figure: 5 Shows that the various feature available through searching.

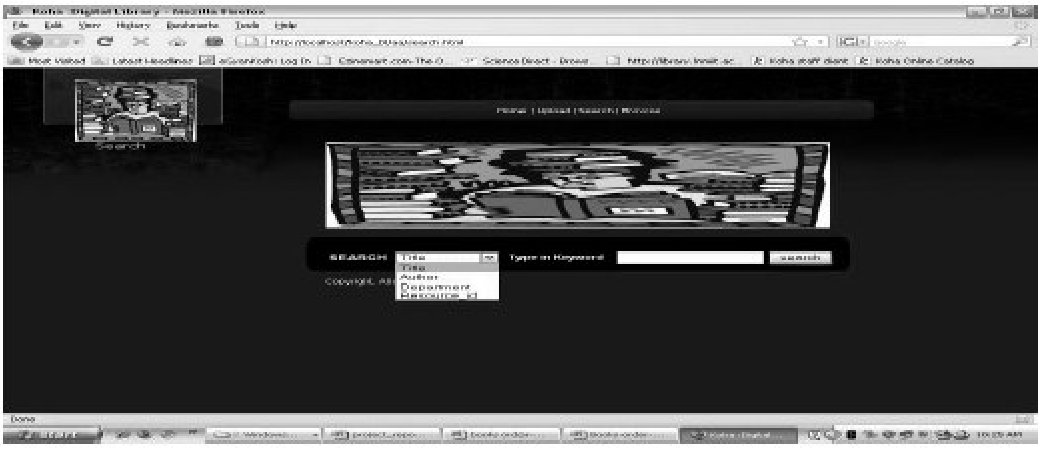

Koha :Digital Library - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Most Visited Latest Headlines eGyanKosh: Log In Ezinemart.com-The O... ScienceDirect - Brows... Koha staff client Koha Online Catalog
Search
Home | Upload | Search | Browse
SEARCH
Title
Title
Author
Department
Resource_id
Type in Keyword
search
Copyright. All
Done
start

Figure: 5 Feature Available Through Searching

4. 4 Browsing

This feature facilitates user to browse the content based on various fields. This interface also facilitates user to browse the content alphabetically. The idea to enable this feature has been taken from the dspace (http:://www.dspace.org). Figure 6 shows the various features available through this module.

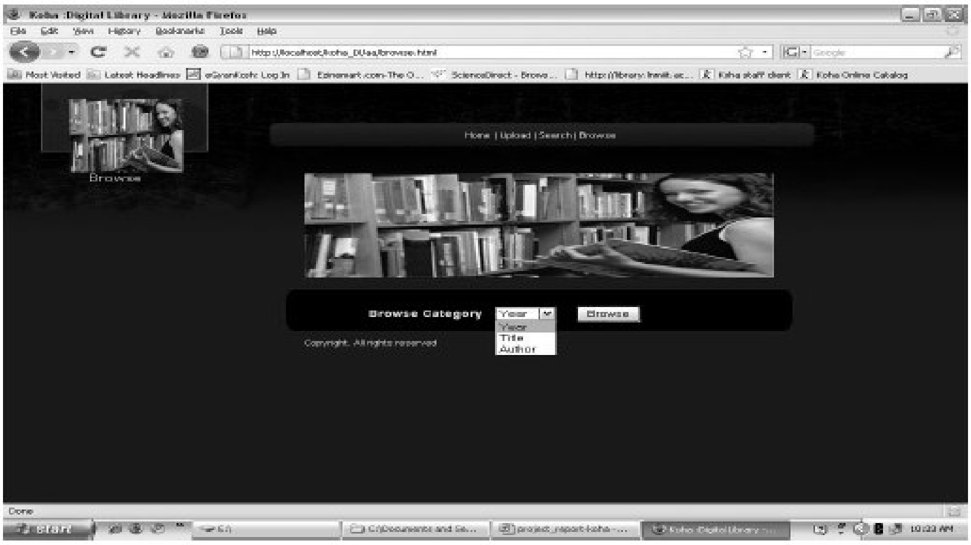

Koha :Digital Library - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://localhost/koha_DUaa/browse.html
Most Visited Latest Headlines eGyanKosh: Log In Ezinemart.com-The O... ScienceDirect - Brows... Koha staff client Koha Online Catalog
Browse
Home | Upload | Search | Browse
Browse Category
Year
Year
Title
Author
Browse
Copyright. All rights reserved
Done
start
10:22 AM

Figure: 6 Feature Available Through Browsing

5. Conclusions

This study sought to understand the prerequisite to understand a basic approach to develop digital library software. koha 3.0 an open source ILS. However, it explore the opportunity to use koha as a DL but as there are two approaches to integrate separate module to koha discuss above that the php scripts should be kept within koha software. The study explores the idea to use koha as a Digital Library but it works on http://localhost/ not within koha software and port: 8080. The study had other limitations, which must be taken into account. It did not explore the features of password change and the administrator approval for the content which are extensively used in many of the digital library software.

This software is a general approach in terms of the programming, data structure and MySQL table layout. One could explore further by using Java and Ajax to make it more interactive and powerful. It is hoped that this DLS will benefit the libraries. However, with more modification and metadata enrichment, and by adding other features of DLS, it will be more useful.

References

1. Cervone, Frank. (2003).The open source option. Library Journal NetConnect(Summer),8-12.
2. Digital Library Federation (1998). A working definition of digital library Available at: www.diglib.org/about/dldefinition.htm
3. E.S. Raymond. Homesteading the Noosphere(1998). First Monday, 3(10)
4. Hars A and S. Ou(2002). Working for Free? – Motivations of Participating in Open Source Projects. International Journal of Electronic Commerce, 6,25–39
5. Koha Developer Wiki. (2007). Koha users around the world.Retrieved, from http://wiki.koha.org/doku.php?id=kohausers
6. Levesque, M.(2004.) Fundamental issues with open source software development. First Monday 9 (4). Retrieved from http://firstmonday.org/issues/issue9_4/levesque/index.html
7. Morgan, Eric Lease.(2002). Possibilities for open source software in libraries. InformationTechnology and Libraries 21 (1):12-15.
8. Netcraft(2005) Web Server Survey Retrieved from URL http://news.netcraft.com/archives/web_server_survey.html
9. Nichols, D. M., Bainbridge, D., Downie, J. S., and Twidale, M. B. (2006). Learning by building digital libraries. Proceedings of the 6th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'06). ACM Press, New York, NY,185-186

About Author

Mr. Shriram Pandey, Assitant Librarian, The LNM Institute of Information Technology, Jaipur (Raj).
Mr. Pankaj Kumar Singh, Assitant Professor, Guru Ghasidas University, Chattishgarh.