Systematic approach of data migration, customization and implementation of Koha: a case study of Saharanpur Campus Library, IIT Roorkee

By

Vinod Kumar Mishra
Assistant Librarian
NIT Rourkela
mishravk79@gmail.com

1. Abstract

Paper describe process for migration of exiting data, implementation and customization of open source software (OSS) Koha, problem faced and discovering their solutions from various sources, to share customized codes and motivate library professionals to use OSS without spending funds with little effort. Paper discusses Saharanpur Campus Library of Indian Institute of Technology Roorkee (IITR) Library automation and data migration processes for Implementation of Koha. It highlights the methods, tips and tricks; configurations, customizations, and migrating of data in new integrated library management system (ILMS) koha. Paper proved that it is not necessary to hire external commercial supports for implementation and migration of data to new OSS Koha. Whole process needs testing and using tips and tricks instead of much technical knowledge on computer and computer language. It also reveals that solutions of most problems have been provided by open source tools and technology which needs to be only discovered. Koha reduced efforts of staffs in routine works of the library and drastically enhanced the user based online services along with web 2.0 based services without investing single penny on adoption of new ILMS Koha. Paper disclosed that technical works for implementation of Koha can be done by Library professionals at zero cost with little efforts. Hence the paper is most significant for those who wish to automate their library at zero cost and migrate on to open source standard like Koha without any fear and finance. The codes given can be used to convert Koha home page as library website i.e. Koha website cum OPAC on the same server.

Keywords: Koha, MarEdit, Winisis, Integrated library management system, Customization, Ubuntu, Library automation, India, Open Source Software, IIT Roorkee

2. Introduction

Paper attempts to discuss about the real practical problems, solutions, technical difficulties and other management related issues while migrating from existing ILMS to an open source solution (KOHA). Koha is the most popular ILMS having advanced user based services and modules for almost all housekeeping related activities. Paper reveals step by step data migration process, configuration of koha system, conversion of data from existing legacy system to standard MARC format and customization of Koha. Various sources and technical requirements also discussed

which help other libraries to implement and migrate data from existing system to Koha. The case study is about the IIT Roorkee, Saharanpur Campus Library which used Troodon3 commercial software which lacks many features and non-standard modules like cataloguing etc. hence it was felt to go for new ILMS for ease of routine work of the library and latest advance feature for the benefit of users.

3. Literature Review & Methodology

Lot of online resources has been consulted by the author during the whole process of data migration, customization and implementation of Koha which is given under reference section of this paper. It was also felt that documents on the required topic are also lacking but online resources and community support is sufficient for the entire work. Methodology is empirical and adopted to provide immediate solutions for the existing problems at Saharanpur Campus Library of IIT Roorkee. An observation and hit and trial methods has been adopted with the consultation with online resources and community support as and when required during the process.

4. About Saharanpur Campus Library

Saharanpur Campus of IIT Roorkee offers academic programs at various levels in Pulp and Paper Technology, Polymer Science and Technology, Process Engineering and Management. The Department is the predecessor of the School of Paper Technology established by the Government of India in 1964, with an aid from the Royal Swedish Government. Saharanpur Campus of IIT Roorkee is having three departments which are recently tri-furcated from Department of Paper Technology, hence the library is having rich collection of Paper Technology related resources in India, apart from above library is also having good collections of books on engineering, management, humanities, Hindi and English literatures etc. Approx 30 thousand collections of library has been divided into General books, reference books, Book bank, reserved collection, Theses, dissertations, B.Tech. Projects, CD/DVD's, and journals. The library serves more than 600 users divided into various groups like Faculty, Under Graduate, Post graduate, Research scholar, non teaching staff etc. with various types of their circulation privileges, approx 50 transactions carried out per day. Library provides all the modern facility like internet access, Xeroxing, computing etc. Library performs all housekeeping tasks like accessioning, cataloguing and other technical processing of books after receiving from the vendors/Central library located at main campus, Roorkee. Unfortunately library strive a lot due to non-filling of vacant post and at present there are two professional staff, one attendant and two Library science trainees but lacking knowledge of computer and expertise in Linux etc. Library runs in two shifts from 0845 to 2000 hrs and open from Monday to Friday.

5. Need for new ILMS

The library was semi-automated with commercial software Troodon3 which was implemented around 2004 with the support from IIT Roorkee main campus central library since it was purchased by them with the agreement to use at any department/campus library of IITR. Troodon3 lacks many features, like setting up of circulation privileges for various type of

documents, fine collection, webopac, compatibility for implementation of radio frequency identification (RFID), web based user services, web2.0 etc. because of which staff members and users were dissatisfied in modern context of library system and services. So library decided to go for the change of the ILMS which has most modern features and could fulfill the user's and staff's requirements. Due to lack of staff it was also necessary to use the modern ILMS which could facilitates easy in working with minimum efforts and staffs.

6. Problem With existing software

There were several problems faced due to existing software for routine works of the library and user based services some of the major problems are:
- Lack of webopac module for browser based searching throughout the campus.
- Lack of good configuring features to set circulation privileges for various type of users and documents.
- Lacking more features and flexibility related to fine management.
- Lack in User friendly and non-standard Cataloguing module.
- Lack of features related to Bar-code and User ID card generations and configurations.
- Apart from above it was lacking all the features in modern context which is described in next sections.

7. Selection of Koha

Mahatma Gandhi Central Library (MGCL) at IIT Roorkee main campus migrated from Troodon software to Libsys so it was decided that same system could be adopted at our library but due to distance of several kilometers in between and some technical issues the decision was dropped. This is how idea for implementing KOHA came into being. Koha is more popular in academic libraries due to its advantages like powerful Circulation, Cataloguing, OPAC and other housekeeping features having advanced web2.0 features through which users can participate in knowledge and library resource sharing's, evaluate and suggest the resources. The most important is it can be used without any restriction of data storage, accessibility and at zero cost.

8. Installation and testing of Koha

The major challenge was to gain reasonable knowledge of LAMP architecture and install KOHA on LINUX. The Live CD/DVD on Ubuntu available on the internet made the task a lot easier. Live CD/DVD iso file was downloaded from the sourceforge website. With many interpersonal communications with the experts and discussions on the open forums Koha was installed independently by author on a 512 MB RAM PC for running a prototype and testing which took almost one month. Several issues like availability of good computer, knowledge of software like Apache, Mysql, Perl language etc., and non availability of technical manpower were confronted in the implementation and configuration part but online community help resolved them too. Some minor corrections on live DVD were also informed to the concern person and discussed for other technical issues and their solutions. Sample data of all type of library materials and users category were created with different circulation privileges than thoroughly tested the

software as per requirements and discovered the solutions of the problem/requirements given beautifully in the Koha system.

9. Data migration

The trickiest part of whole process was migration of data from the existing system to the new one. The existing ILMS database was based on MS_Access and there was no database password to see the structure of data inside MS_Access, so discovered option in Troodon3 to export the whole data (Library materials except user data) in ISO 2709 format. Exported data was not sufficient because conversion of ISO 2709 into Marc format was another challenge but marc editor made the job easy which can convert the CDS/ISIS iso format (Winisis) database to Marc format as discussed in next section.

9.1 Application of Winisis (CDS/ISIS) in data migration

The author imported the data in CDS-ISIS software which was got through Troodon import option in iso 2709. After successful import and random checking of data in CDS-ISIS again data was exported from CDS-ISIS in iso format and converted into marc format with the help of "CDS-ISIS.iso to MARC Translation" option given in marc editor software.

Fig.1: Structure of Original iso-2709 database in Winisis
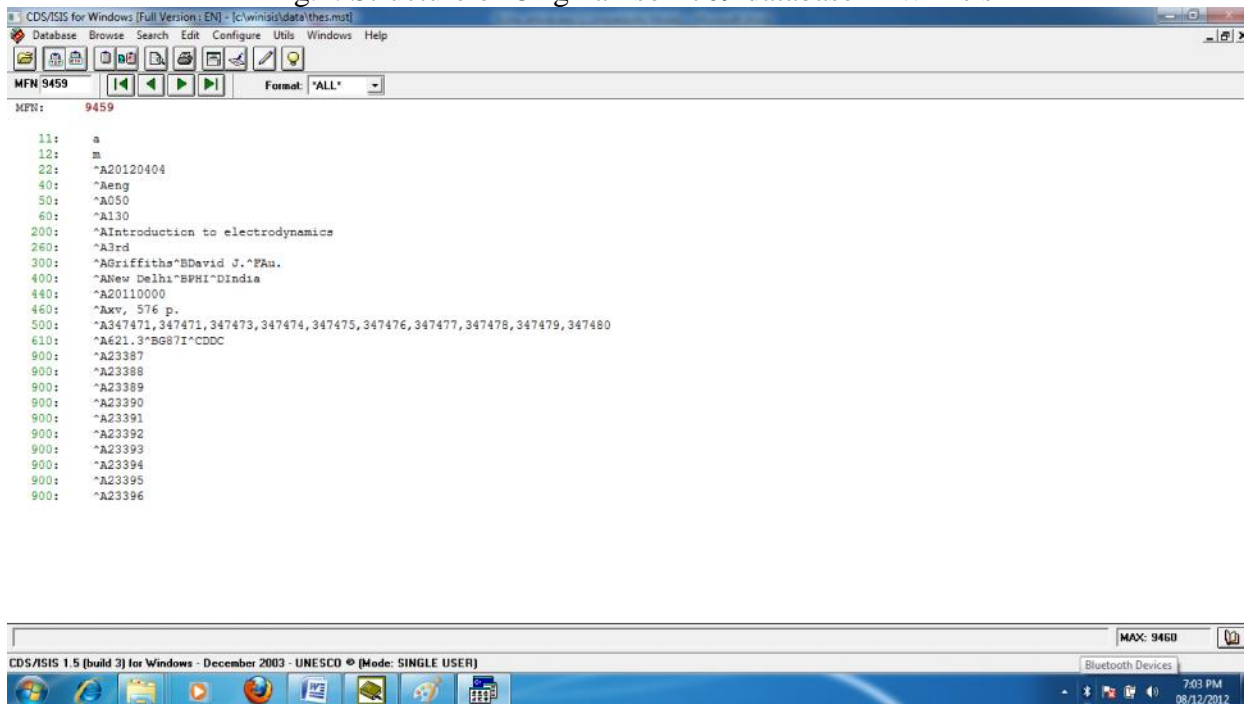
9.2 Marc editor
Marc editor is user friendly solution to convert CDS-ISIS iso data in Marc format. Author found that converted data from CDS-ISIS to marc in above step was completely nonstandard (an example given below) hence lot of standardization was needed to import it in Koha ILMS.

```
=LDR  01008000000003610004500
=011  a
=012  m
=020  $AIPT
=022  $A20011019
=040  $Aeng
=050  $A010
=060  $A100
=200  $AEncyclopedia of chemical technology
=260  $A4th
=300  $AOthmer$BKirk
=400  $ANew York$BJohn Wiley
=440  $A19980000
=460  $AV 25- xxviii,912 V 8- xxviii,1093 V 23- xxviii,1118 V 14- xxviii,1115 V 18-
xxviii,1073 V 21-24 148
=490  $AVol 25$CVitamins to zone refining
=490  $AVol 8$CDeuterium and tritium to elastomers, polyethers
=490  $AVol 23$CSugar to thin films
=490  $AVol 14$CImaging technology to lanthanides
=490  $AVol 18$CPaper to pigment dispersions
=490  $AVol 21-24$CIndex Recycling, oil to vinyl polymers
=500  $AIPT
=610  $A660.03$BK73$CDDC
=620  $AChemical technology- Encyclopedias
=900  $A17213$B25$CR
=900  $A17372$B8$CR
=900  $A17052$B23$CR
=900  $A17566$B14$CR
=900  $A16897$B18$CR
=900  $A17212$B21-24$CR
```

Therefore mental analysis and searching the process on how to make it standard helped in finding several useful files over the internet but all were just to convert good data into standard marc. So systematic approach to achieve the standardization and most satisfaction are discovered and explained in next section.

9.3 *Mapping and editing of bibliographic data in correct MARC format*

Comparison of different tags of Troodon and marc tags of Koha from Library of congress (LOC) website was done and suitable standard tags were found as per requirements of Koha.

Table.1: The various tags in Troodon and their equivalent in Koha/Marc

Description	Troodon Tags	Marc/Koha Tags
Leader	=LDR	LDR
	=011	
	=012	

Source of Record	=020	037 $b (May be in 952$e)
	=022	Need to remove
Language	=040	041 $a
Physical Medium	=050	
Type of material	=060	942 $a (Delete this number and insert fresh field)
ISBN	=100	020 $a
CODEN	=102	
Title	=200	245 $a
Journal Name	=201	
Edition	=260	250 $a
Name of persons	=300	100 $a ($F Au,comp. etc need to be removed)
Corporate body	=310	110 $a
Name of Meeting	=320	111 $a
Publisher	=400	260 $aPlace$bPublicher$cYear
Date of publication	=440	260 $c
Serial Number/Date	=450	
Physical Description	=460	300 $a
Price	=465	
Series statements	=480	490 $a
Part statement	=490	505 $a
Note	=500	500 $a
Bibliographical relation	=510	
Frequency	=520	
Call Number	=610	082 $aClass$bBook No.
Descriptors/Subject heading	=620	650 $a
Bill No./Date/Vendor	=800	
Accession number	=900	952 $p

Additional codes as per Koha requirements for various reading materials like reference Books; General Books, Book Bank etc. was added in Marc format as shown in table 2.

Table 2: item types and their codes used in Koha for data conversion

Item Type	Code
Books	BK
Reference Books	REF
Book Bank	BB
Reserve Books	RB
Location of Items	**Code**
General Section	GEN
Reference section	REF
TBLS/Books book section	TBLS
Reserve section	RS

The mapping was done only for the tags under which bibliographical data of library materials were present in the Troodon database as shown in table 1. Therefore fully exploited the options of Marc editor software and converted all basic Troodon tags one by one into standard Marc tags. Inserting, editing, replacing, deleting etc. related work was more complex but some of the things has been done with the options available in Marc editor and one by one all complex nonstandard data was successfully converted into standard Marc format with tricks. The final converted data of Troodon as shown in section 9.2 above now looks like below as per requirements of Koha software:

=LDR 01665nam a2200313Ia 4500
=008 101213s1998 utu o eng d
=037 \\$bIPT
=041 \\$aeng
=082 \\$a660.03$bK73
=100 \\$aOthmer, Kirk
=245 \\$aEncyclopedia of chemical technology
=250 \\$a4th
=260 \\$aNew York:$bJohn Wiley,$c1998
=300 \\$aV 25- xxviii,912 V 8- xxviii,1093 V 23- xxviii,1118 V 14- xxviii,1115 V 18-xxviii,1073 V 21-24 148
=505 \\$aVol 14. - Imaging technology to lanthanides
=505 \\$aVol 18. - Paper to pigment dispersions
=505 \\$aVol 21-24. - Index Recycling, oil to vinyl polymers
=505 \\$aVol 23. - Sugar to thin films
=505 \\$aVol 25. - RBitamins to zone refining
=505 \\$aVol 8. - Deuterium and tritium to elastomers, polyethers
=590 \\$aIPT
=650 \\$aChemical technology- Encyclopedias
=942 \\$cREF
=999 \\$c3003$d3003
=952 \\$w2012-05-02$p16897$r2012-05-02$40$00$6660_030000000000_K73$95167$bMAIN$10$o660.03 K73$d2012-05-02$t18$70$cREF$yREF$aMAIN
=952 \\$w2012-05-02$p17052$r2012-05-02$40$00$6660_030000000000_K73$95168$bMAIN$10$o660.03 K73$d2012-05-02$t23$70$cREF$yREF$aMAIN
=952 \\$w2012-05-02$p17212$r2012-05-02$40$00$6660_030000000000_K73$95169$bMAIN$10$o660.03 K73$d2012-05-02$t21-24$70$cREF$yREF$aMAIN
=952 \\$w2012-05-02$p17213$r2012-05-02$40$00$6660_030000000000_K73$95170$bMAIN$10$o660.03 K73$d2012-05-02$t25$70$cREF$yREF$aMAIN
=952 \\$w2012-05-02$p17372$r2012-05-02$40$00$6660_030000000000_K73$95171$bMAIN$10$o660.03 K73$d2012-05-02$t8$70$cREF$yREF$aMAIN

=952 \\$w2012-05-02$p17566$r2012-05-
02$40$00$6660_030000000000000_K73$95172$bMAIN$10$o660.03 K73$d2012-05-
02$t14$70$cREF$yREF$aMAIN

9.4 Configuration of Koha

Configuring koha system as per Marc data was also done except user setup like circulation privileges, category etc. like codes for library (MAIN), authorized value for location of different item type (Reserve section, Reference section etc,) and various codes for Item types etc.
The Marc bibliographical framework for Book as shown in table 3 for different type of library book was configured and setup to import all the converted marc data in Koha.

Table 3: Marc bibliographic framework for Book

Tag	Description
000	**LEADER**
00	fixed length control field
008	**FIXED-LENGTH DATA ELEMENTS--GENERAL INFORMATION**
00	fixed length control field
020	**INTERNATIONAL STANDARD BOOK NUMBER**
A	ISBN
037	**SOURCE OF ACQUISITION**
A	Source of stock number/acquisition
041	**LANGUAGE CODE**
A	Language code of text/sound track or separate title
082	**DEWEY DECIMAL CLASSIFICATION NUMBER**
A	Classification number
B	Item number
100	**MAIN ENTRY--AUTHOR NAME**
A	Personal name
110	**MAIN ENTRY--CORPORATE NAME**
A	Corporate name or jurisdiction name as entry element
111	**MAIN ENTRY--MEETING NAME**
A	Meeting name or jurisdiction name as entry element
C	Location of meeting
D	Date of meeting
245	**TITLE STATEMENT**
A	Title
B	Remainder of title
C	Statement of responsibility, etc
250	**EDITION STATEMENT**
A	Edition statement
260	**PUBLICATION, DISTRIBUTION, ETC. (IMPRINT)**
A	Place of publication
B	Name of publisher
C	Year of publication
300	**PHYSICAL DESCRIPTION**

A	Number of Pages
B	Other physical details
E	Accompanying material
440	**SERIES STATEMENT/ADDED ENTRY--TITLE**
A	Title
490	**SERIES STATEMENT**
A	Series statement
500	**GENERAL NOTE**
A	General note
505	**FORMATTED CONTENTS NOTE**
A	Formatted contents note
520	**SUMMARY, ETC.**
A	Summary, etc
650	**SUBJECT ADDED ENTRY--TOPICAL TERM**
A	Topical Term
X	General subdivision
700	**ADDED ENTRY--PERSONAL NAME**
A	Personal name
856	**ELECTRONIC LOCATION AND ACCESS**
U	Uniform Resource Identifier
942	**ADDED ENTRY ELEMENTS (KOHA)**
C	Koha [default] item type *

Apart from above bibliographic framework some field for individual items (Tag 952 in Marc) as shown in screenshot (fig.2) was also selected which appears after feeding all the bibliographic details of the items in bibliographic framework for Books.

Fig. 2: Marc tag 952 selected for each individual item

9.5 Importing/exporting of Marc data and re-editing the data in Koha ILMS

After successful import of data in Koha through the option **Stage MARC records for Import** under **Tools** option we checked the data in real OPAC and found some errors, than again corrected and re-imported the data in koha, the process of re-importing and editing of Marc record were done several times until achieved the maximum satisfaction/accuracy. Author found several duplicate data of the records due to poor data entry quality in previous software but successfully merged with option of koha which reduced unnecessary size of database, circulation related issue and reporting related problems.

9.6 Import of Users Data in Koha

User data was separately imported with no relationship with book issued to members, this was done due to lot of corrections were needed to be done in Books database like changing of accession number which was non-uniformly entered in the Troodon database. Marc editing was suitable stage for corrections of all the errors which made the job easy by tricks of marc conversion. Since maximum books were returned by users during summer vacation so we did not faced any circulation related issues hence circulation data was not imported in Koha. Only member's data copied from the Ms_access database of Troodon and converted in .csv format as instructions provided in downloadable sample format in Koha staff client.

10. Customization of Koha Search and Display

After successful migration/correction of all Book and users data felt need of some additional search and display features for ease of work and user comfort, like browsing of whole item type, searching by acc. number in basic search and display of accession number in detailed display for ease of placing hold on specific item. Hence all were done successfully as discussed below.

10.1 Customization for Catalogue browse & search by accession number

For browsing, discovered tip from the koha OPAC available on various koha websites and simply copied the code inside our Koha system preferences and changed the Items code which was assigned for our library item types (Fig.3).
Searching by accession number is very frequently needed mainly by staffs for some technical reasons but lack of this feature on main koha home page bound us to use advance search by multiple clicks which takes much time to reach up to that option. Therefore the code in file **masthead.inc** was carefully observed and modified to get restricted search by accession number in koha main home page.

10.2 Customization for display of accession number

Other important feature related to display of accession number in detailed display was missing due to which patrons and staff had to click another option i.e. Marc view. Therefore it was necessary to add this feature which facilities ease in placing hold on specific item by patron and several data correction related work to library staffs. The code in **opac-detail.tt** file is

responsible for display of some book related information like class number; location etc so to get the above effect new code was inserted.

11. Koha Home page customization

After concluding all the above works successfully author found possibility to customize the koha home page in such a way that it could serves the users as Library website cum online public access catalogue. This idea was wonderful to have a website and OPAC on single server and on single web.2.0 based ILMS koha, but it was more challenging due to lack of expertise in HTML, CSS etc. Keeping in view the importance of library website we made list of links and contents which may be needed to display over the website and the possibility of putting that idea over the Koha OPAC home page. The whole things were discovered from various internet resources and some basic book on HTML and CSS which is described in sub-sections below.

11.1 Display of drop down menu

Creating drop down menu over the webpage is good to have more linking in minimum space of the page so author discovered the code which is available free of cost at http://www.spiderwebmastertools.com and modified as per library requirements after consulting some basic books on above requirements, the code which is used to display the drop down menu as shown in fig. 03.

Fig.03: Home page shows drop down menu in pink color

11.2 *Display of Left navigation menu, Credits area & Other customization*

The left navigation of system preference used to provide most important frequently used informative links regarding help and updated information. Credit code used to display copyright and contact information is beautifully placed in the credits option of system preference.

Apart from above customization, color of the home page changed by modification of color code in opac.css file, the logo of koha is replaces by Indian national flag and the header image is beautifully constructed with the help of paint tool of windows. Several html and doc files related to information about library were placed in the newly created folder and linked through drop down menu which worked well. Facebook liking and send link code was also placed in credit area to popularize the library website cum OPAC. To check the popularity, traffic and other user access related behavior of Koha website cum webopac author also placed the free website counter/monitor code.

12. User perceptions

Since many user based features were lacking in the Troodon3 software like OPAC, online user based services like reservation, renewal etc. the library fine statistics shows that the collection of overdue charges decreased to half of the past years statistics due to lack of online renewal facility in Troodon3. Online reservation facility made users more satisfied due to getting their required books when needed, Uses of resources of library maximized as revealed by the statistics of Circulation and various reports which was lacking in previous software. Efforts, staff interference and time consumed in various housekeeping activities much reduced due to standard features in koha like Z39.5 feature for cataloguing of library books from Library of congress websites and many others. Online renewal facility helped in reducing time and manpower at circulation desk too. Users enjoyed sharing of their reading list and interested books on social network sited as well as star rating of the popular books. Option of Most popular reading helped library for the acquisitions of high demanded books. Web based features of Koha for staff and users helped them in getting uninterrupted services on their Desktop/Laptop without any technical issues and hardware/software requirements round the clock.

13. Conclusion

The paper very well described the whole process of implementation of koha at IIT Roorkee. The case study described step by step work of data migration successfully with the help of solutions available on web resources. The koha software facilitates many advance online features like renewal, reservation, comment on books, along with strong features of Cataloguing, Circulation, Acquisition and Serial control etc. The most important is all the process involved zero cost i.e. data migration, customization and final implementation of koha successfully with most advanced features hence it is most suitable solution for libraries whose main concern is finance. Koha provides fully web 2.0 based facilities hence the OPAC home page can be easily converted into Library website. The paper reveals availability of online resources which can be used by library professionals for Library automation with Koha, data migration, and designing of websites.

14. References

a. Department of paper technology, IIT Roorkee available at: http://www.iitr.ac.in/departments/DPT/pages/About_Us+History.html (accessed on 11 November 2012).

b. Spider web master tools, available at: http://www.spiderwebmastertools.com (accessed on 30 June 2012).

c. Koha, available at: http://koha-community.org (accessed on 02 April 2012).

d. Mysore University library OPAC, available at: http://libcat.mysore-univ.org/ (accessed on 06 June 2012).

e. Marc standards, available at: http://www.loc.gov/marc/ (accessed on 01-30 May 2012).

f. MarcEdit, available at: http://people.oregonstate.edu/~reeset/marcedit/html/index.php (accessed on 01-30 May 2012).

g. Basic Ubuntu commands, available at: http://freshtutorial.com/basic-ubuntu-command-tutorial-for-beginners/ (accessed on 05 May 2012).

h. Koha Live DVD, available at: http://sourceforge.net/ (accessed on 11 November 2012)

i. w3schools.com: the world largest web development site, available at: http://www.w3schools.com/ (accessed on 02-30 June 2012).

j. Kohageek, available at: http://kohageek.pbworks.com/w/page/28248118/FrontPage (accessed on April-June 2012).

k. Koha wiki, available at: http://wiki.koha-community.org/wiki/Main_Page (accessed on April-June 2012).

l. Koha: Customized live DVD, available at: http://sourceforge.net/project/vkmkoha/files/ (accessed on 18-31 January 2013).