Bibliographic Data Migration in KOHA 3.18 from Existing Data Sheet: A Practical Experience

Article · March 2017

2 authors, including:

Sanjay Karak
Netaji Mahavidyalaya
17 PUBLICATIONS   **8** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Data migration to Library automation software View project

Bibliographic Data Migration in KOHA 3.18 from Existing Data Sheet: A Practical Experience

Sanjay Karak
Librarian
Netaji Mahavidyalaya
Arambagh, Hooghly
West Bengal, India
karaksanjay@hotmail.com

Kuheli Dutta
Assistant Librarian
Presidency University
Presidency University
kuheli.lib@presiuniv.ac.in

Abstract

Many academic institutions used library management software for the access, dissemination, store and use of information or store of bibliographic record or store member data, administrative works. These libraries used different types of automation software as like SOUL, Libsys, Koha, e-grantahalaya etc. In this paper I tried to explain the data migration from excel sheet to library automation software database. Our college library ware stored a huge numbers of bibliographic data on excel sheet. I tried to convert theses data to our database and reduced effort, time and monotony of jobs.

Key Word: Koha, MARC Editor, Mata Data Conversation

Introduction:

The application of library automation software change every part of libraries functions and move to the libraries efficient, reliable, cost-effective, time saving. The library automation is to facilities timely accessed of proper information with speed and accuracy. The demands of proper documents and information are increased from users with the help of library automation systems.

Library automation:

Library automation refers to the process of mechanization of library activities, services and routine works such as acquisitions, circulations, administrations, serials control, cataloguing and

OPAC etc. An automated library is one where one and more computer systems are used to manage one and several of the library's functions such as --------

1) The explosion of information.
2) Information in machine readable form.
3) The feasibility of economics.
4) Feasibility for routine jobs.
5) The collections arrangement
6) Avoid the duplication
7) Time saving
8) Speed and accuracy
9) Electronically share the resources.

Koha:

Fully-featured open-source library automation software is Koha. Koha is developed in New Zealand by Katipo communication Ltd and first deployed in January of 2000 for Horowhenua library trust. Koha feature set continues to expand and evolve to meet the needs of its user base.

- *Featured-full LIS:* - Koha is worldwide used software for all types of libraries with comprehensive functionality including basic and advanced options. The modules of koha are to acquisitions, circulations, cataloguing, authorities, serials managements, flexible reporting, multi-format notices, label printing, offline circulation and much more.
- *Translatable and multilingual:* - The large numbers of languages in koha has available and with more languages added every year.
- *Full text searching:* - Koha provides the powerful searching and enhanced catalogue display that can used content from Amazon, google, Open Library, Library thing and much more.
- *Library standers compatible:* - Koha is supported different types of library standers and protocols as like MARC21, UNIMARC, Z39.50, SIP2, SRU/SW, SIP/NCIP etc.
- *Web based interfaces:* - koha's OPAC, circulation management and self checkout interfaces are all base on standards compliant www technologies –XHTML,CSS, and Java script etc.
- *Free software/open source:-* Koha is distributed under the software General Public License (GPL)

Hardware and Software requirements:

A table for compare of minimum hardware and software are given below

1. Hardware compare

Sl. No.		Koha 3.18 requirement	Available in College
1.	Processor type:	Pentium-4, 2.6 GHz or Higher	Intel® Xenon®
2.	Processor speed:	1 GHz or Higher	310 GHz
3.	RAM:	2GB	4GB
4.	Free hard disk space:	80 GB	1000GB

Fig.1

2. Other software

Sl. No.	Windows base	Linux Base
1.	Microsoft Excel/Open office/ Libra office	Microsoft/ Open source
2.	MarcEditor	

Fig. 2

About MARC Editor:

MarcEdit is a metadata editing software suite used primarily to create and manipulate MARC records. Originally developed by Terry Reese in 1999 for a major database cleanup project at Oregon State University, the software was subsequently released for wider use in the Library and Information Science field.

Today, ILS systems are fragmented regarding the type of character set that they will support

1. Two primary character sets:
   *MARC8 (ANSEL) legacy
   *UTF8
2. Most vendors send records in one format or the other, meaning that character conversions are sometimes necessary

Conversion methods:

Our library stores the bibliographic data of 40000 documents in MS-excel sheet. The bibliographic data is stored unscientifically. Finally a map is design for these data to convert in our library database.

The stapes is as like-

1) Shorting the excel sheet data.
2) Rectifying the errors in data.
3) Correcting the excel sheet data.
4) Matching the field and subfield columns of excel sheet as per koha 3.0 software.
5) Transferring the data into the MARCEdito.
6) Converting excel to .MRC format.
7) Converting .MRC to .MRK format.
8) Exporting .MRK data to Koha 3.18 LIS software.

Firstly to start conversion let have an example of raw excel data sheet which is kept in our library.

Fig. 3

In the screen shot of Fig. 4 above defines deferent filed as like –

- Column A: Date of Publication
- Column B: Subject
- Column C: Book title
- Column D: First Name of Author
- Column E: Middle Name of Author
- Column F:Last Name of Author
- Column G: Edition
- Column H: Publisher
- Column I: Year of the Publication
- Column J: Class Number
- Column K: Pagination
- Column L: Accession Number

- Column M: Price
- Column N: Date of entry in the excel sheet

Shorting the excel sheet data: Excel sheet is shorted on the bases of subject to accumulate all the titles of the same subjects at the same place. First Select Column B & Column C in excel sheet and press the button sort & filter.

Rectifying the errors in data: after the shorting all subject and titles next job is to

I.) The double entry title is to move on trace.

II.) Spelling of the titles.

III.) The typographical and conceptual errors of the columns. Example as like "Eco. And Com. Geography" is to be "Economic and commercial geography"

IV.) Error in Indic language documents in English.

V.) Grammatical error.

VI.) Conceptual errors in author field. Example as like the excel sheet column D, E and F is allocated for First name, Middle name and Last name of author. A book is written by "Amit Kumar Das" and sheet is allocated as D= Amit, E=Kumar F= Das, if A book is written by three authors that time the sheet is allocated as D= Amit Kumar Das, E=Krishna Das, F= Kishor Das so for conversion data is corrected one by one.

Correcting the excel sheet data:

After the rectifying about 40,000 rows of excel sheet data. These data is too hard to sort out from excel sheet, so excel formula is applied for combining the texts of two or more particular cells, texts before the authors, moving the error words, expand the short form of words etc. To combine the text from multiple cells to a one cell, ampersand operator used for First name of author and last name of author.

1. Formula for two cells in one cell is = (type&" "&) [a space include in quotation marks]

2. Formula for more than two cells in one cell is= Selecting cells, type &" "& after each cell, then select the next cell. For add a space between combined text type & instead of &" "& to insert a comma, type &","& press Enter.

Matching the field and subfield columns of excel sheet as per koha 3.0 software

Koha, library automation software is compatible with MARC tags. For cataloguing of documents in koha, needs to map which sub fields is essentials and which fields have in our excel sheet. Now let's see

Field name	MARC tag	Excel Sheet collumns
Title	245$a	C
Remainder of the title	245$b	C
Statement of responsibility	245$c	D,E,F
Edition	250$a	G
Author	100$a	D,E,F
Publishing place	260$a	
Publisher Name	260$b	H
Publishing date	260$c	I
Class Number	082$a	J
Pagination	300$a	K
ISSN	022$a	

Fig. 4

Transferring the data into the MARCEditor

Open MarcEdit 6.0 and select "Export tab delimited Text" icon. Then browse excel sheet on source file and output file destination, specify output file name, excel sheet number etc. And also Specify LDR/008 as book.
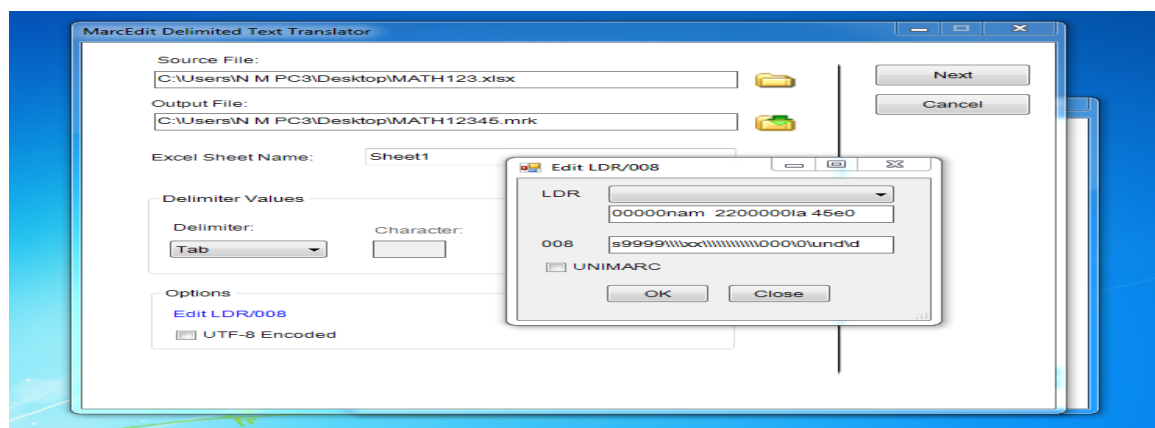
Fig. 5

Converting excel to .MRC format

Most of the library automation software accept .MRK extension file format. So our excel sheet is to convert as an .mrk. After completing field LRD/008 next job is to map excel sheet column to MARC field and sub field. For example field 0 that is fist column I map as 022$a and then click on apply. You can also customize indicator, term, Punctuation etc. After all subfield mapping then clicking on finish button will bring next screen that Marc Text File (.mrc) has been created.
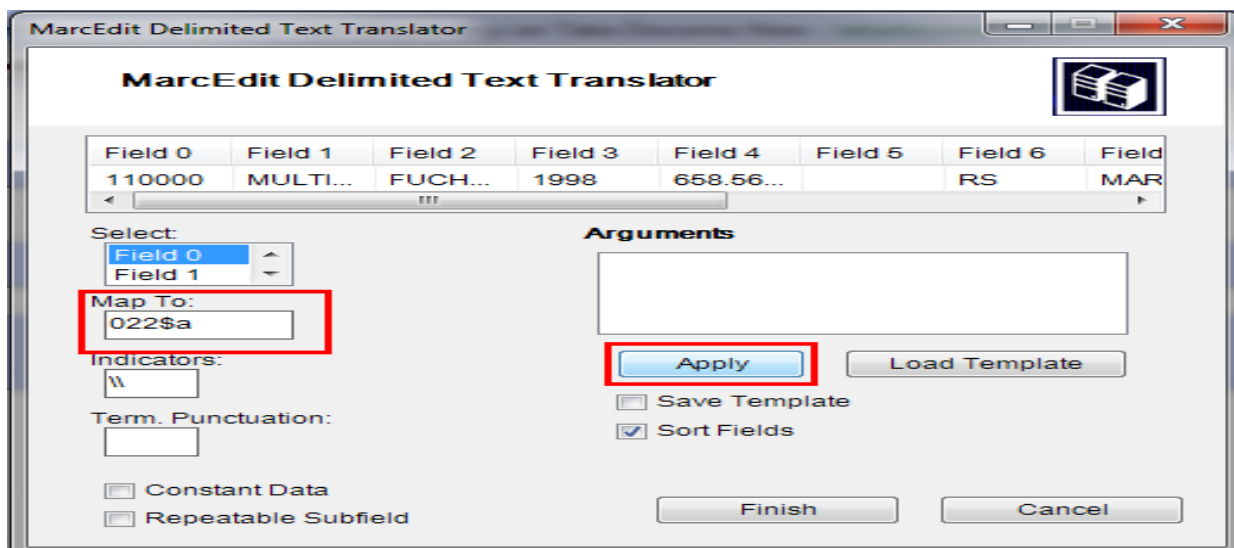
Fig. 6

Converting .MRC to .MRK format

After completing .MRC file, reconvert .mrc to .mrk file. For these reopen MarcEditor and selected MARC tools then browsing input file and given output file name. Next select MarcMaker from functions then click execute, it will show result.
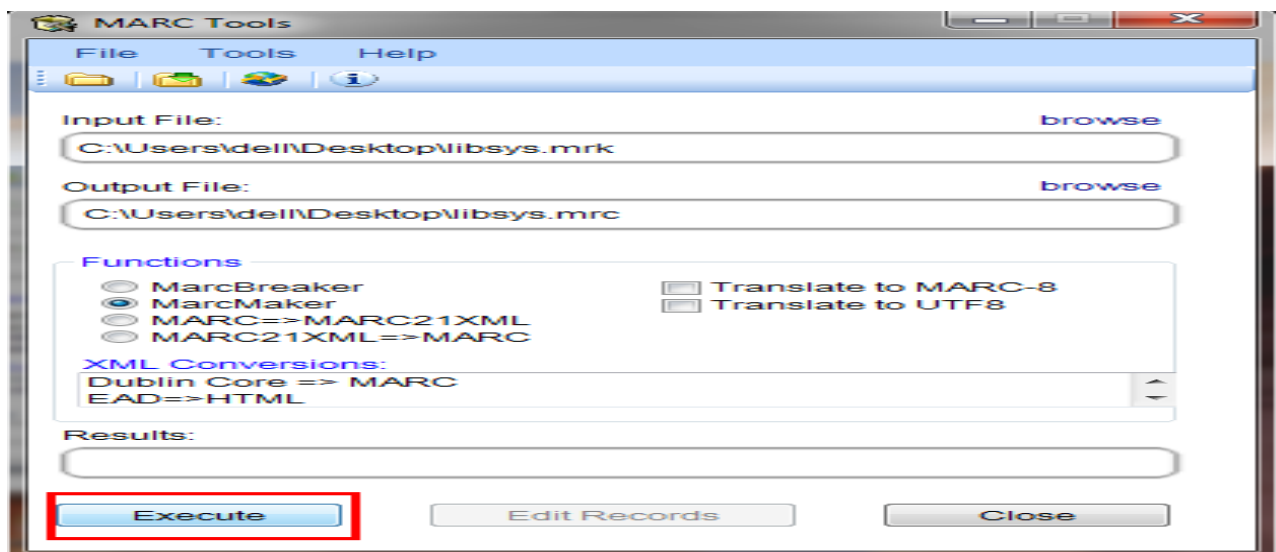

Fig. 7

Exporting .MRK data to Koha 3.18 LIS software

Now .mrk date is ready to move in Koha LIS software. Get there: More>tools>Catalog>stage MARC records for import. This tool can be used to import both bibliographic and authority records that are saved in MARC format. Importing records into koha include two stapes. First is

to stage records for import, Next options for matching and imports. So browse file to stage records into the reservoir.
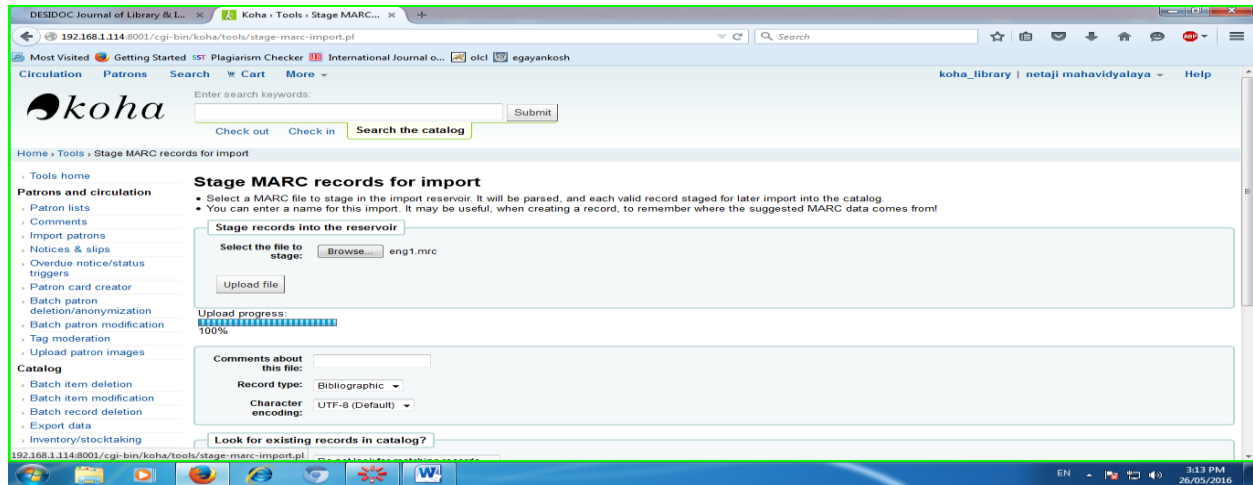
Fig. 8

Then upload file. Enter 'Comments about file' to identify your upload when going to manage staged MARC record tool. Tell koha which type of this is, bibliographic or authority
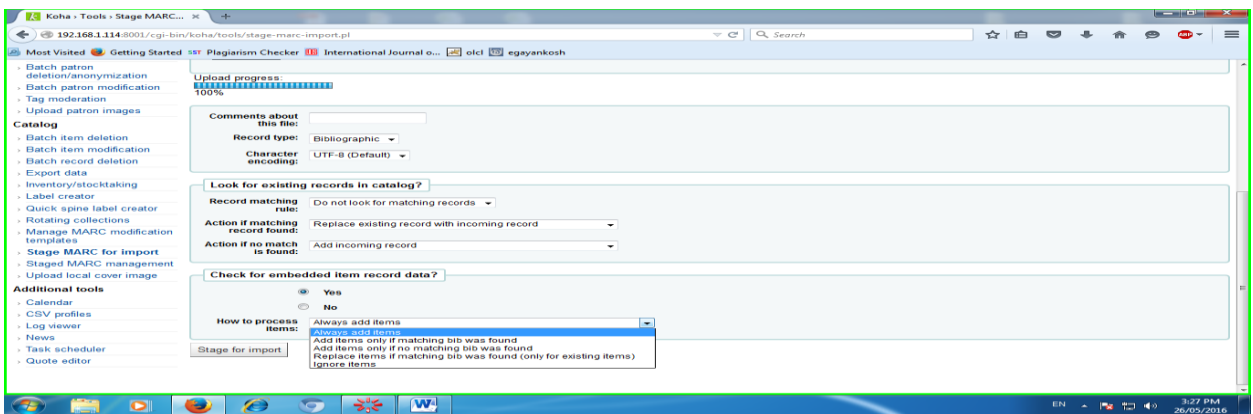
Fig. 9

Next choose whether or not to import the item data found in the Marc record. Form here you can choose to always add items regardless of matching status, add them only if matching bib was found, add items only if there was no matching bib record, replace item if matching bib was found or ignore items and not add them.
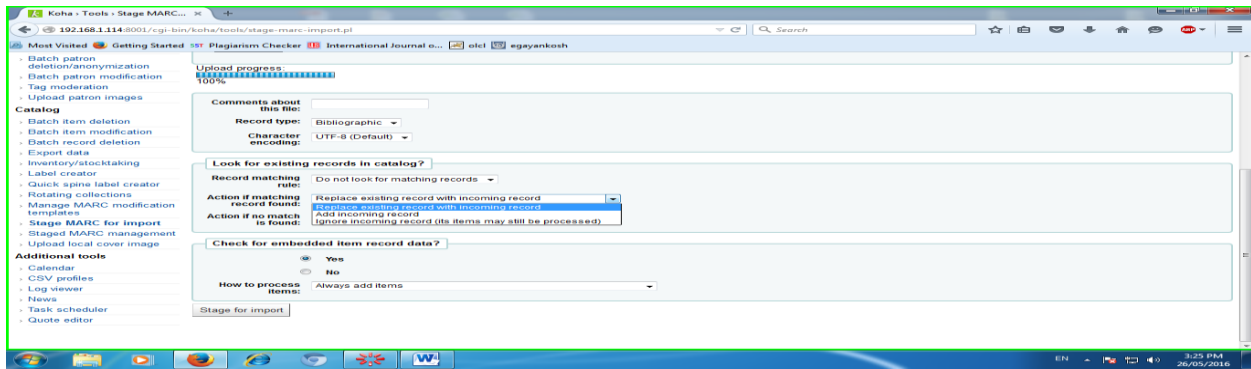
Fig. 10
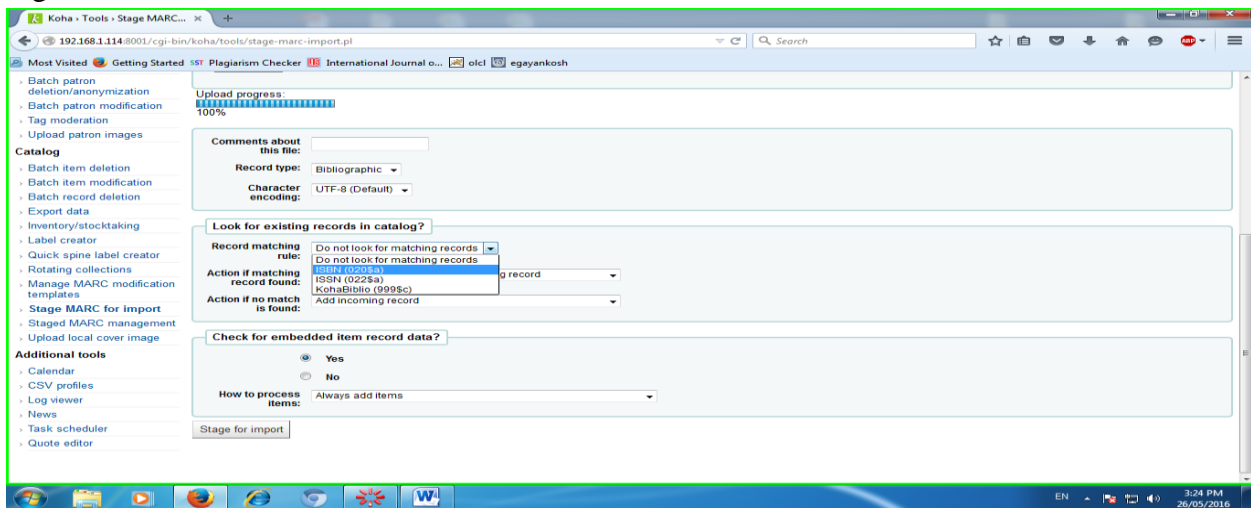
Fig. 11
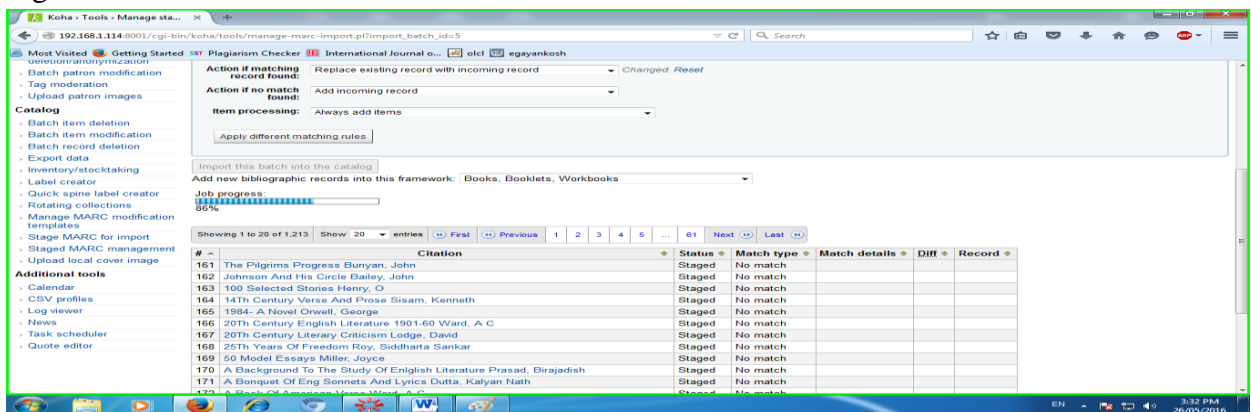
Fig. 12

Marc import is completed.

Fig. 13

Conclusion

Now a day most of the manual libraries are moved as automated library systems. And these libraries are maintenances their record on the accession register or in excel sheet of computer. So, if any libraries moved to automation systems they need huge amount data fed to data base to establish these systems. For these feed need lot of man power and money. We reduce man power, time and money to apply this method. So, automation systems are very important in the libraries. Time saving, avoid of duplication, ordering, time and many more functions are done through Koha 3.18 Library automation software. Acquisition reports are needed to document performance statistics and summaries of work done during the acquisition process in a specified time frame. In the past, they were compiled manually, a tedious job. Nowadays library systems come with a graphical report interface that enables point and click selection of appropriate date elements, time periods, frequencies, output devices etc.

Bibliography

1. Haravu, L. J. (2009). Automated Acquisition System. In M. K. Rao (Ed.), *Library Automation.*
2. Mohsin, Syed Fayyaz(2013). Comparative Study of Library Automation Software (Special Reference of SLIM21 and SOUL)
3. http://marcedit.reeset.net/about-marcedit (Accessed on 24.08.2015)

4. Arora, J. (2009). State of the Art Library Automation System, Ahmedabad: INFLIBNET Cente
5. Reese, Terry (14 March 2013). *"About MarcEdit"*. MarcEdit Development. *Retrieved 22 August 2016*.
6. http://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/17350/marcedit_discussion.pdf(Accessed on 28.08.2015)
7. http://www.koha.org/ (Accessed on 29.08.2015)

Follow us on: IRJLIS, Facebook, Twitter