ISSN (Print): 0972-2467
ISSN (Online): 0976-2477
SRELS Journal of Information Management, Vol 56(2), April 2019, p. 85-90 DOI: 10.17821/srels/2019/v56i2/141234

Data Migration from Libsys to Koha ILS: An Experimental Study

Sambhunath Sahoo* and Mukesh Saikia

Central Library, Tezpur University, Tezpur – 784028, Assam, India; sambhumlis@gmail.com

Abstract

Library automation is one of the primary activities of all kinds of libraries. In this paper, we are focusing on academic library system. For library automation, academic libraries are using some Integrated Library Management System (ILMS) software that may be proprietary or open source. The question of data migration emerges when library administrators think of moving from an existing ILMS to a new ILMS. This paper aims to present a systematic approach for data migration from existing ILMS (Libsys) to a new ILMS (Koha) and it has been successfully experimented at Tezpur University library system. The study found that both bibliographical data and patron data in the new system (Koha) are identical with existing system (Libsys).

Keywords: Data Extraction, Data Validation, Data Migration, Data Transformation, ILMS, Koha, Library Automation, Libsys

1. Introduction

Academic libraries all over the world are using some Integrated Library Management System (ILMS) software that may be proprietary or open source for library automation. For a variety of reasons, library administrators may think of migrating from an existing ILMS to a new ILMS. Some of the major factors that suggest the need to migrate could be the problems faced by the library administrators such as price constraints, lack of advanced features in existing system, lack of continuous revisions and updates in the existing system, etc.

Data migration is the process of transferring data from existing system to a new system. New system may not accept data directly from any existing system as the architecture of the new system may be different from that of the existing system. Hence, it needs data transformation. Data transformation is the process of transforming data from one format (source system) to another format as desired by the new system (targetsystem). In this paper, we are discussing data migration from Libsys, which is a proprietary ILMS software to Koha, which is an open source ILMS software. This empirical study deals with data extraction, data transformation and data migration from Libsys to Koha.

2. Objectives

The main objectives of the study are:

- Extracting various kinds of data from existing ILMS (Libsys)
- Checking the status of extracted data, removal of redundant data, and data rectification
- Transforming the extracted data into the format desired by new ILMS (Koha)
- Migrating the transformed data to new ILMS (Koha)
- Verifying migrated data in both the ILMS

3. Literature Review

Data migration from one system to another system is a cumbersome task. Some papers and online resources on the subject have been reviewed in the following paragraphs to get some idea about data migration from Libsys to Koha.

Egunjobi and Awoyemi (2012) shared an experience of library automation using Koha and suggested that library automation requires proper planning and technical support. Avery (2016) conducted a case study and described the implementation of an open source ILS Koha by migrating from a proprietary ILS Destiny

*Author for correspondence

in a special focus institution. He also demonstrated how libraries can migrate efficiently without any vendor support. Pund and Jain (2016) in their study explained the transformation and migration process from proprietary ILS Libsys to open source ILS Koha. Todd (2018) shared an experience of data migration from Millennium to Koha accomplished by a dedicated team of librarians and successfully transferred more than 48,000 records in less than two months. Bissels (2008) discussed the procedures of data migration from proprietary ILS to Koha ILS with the support of a vendor at CAMLIS, Royal London Homoeopathic Hospital, UK.

House suggests that Koha is an open source ILS with huge online support, easy to install and manage (House, 2016). Muller (2011) conducted a study and described how to choose a free and open source ILS. The study found that Koha ILS is the only software that meets several criteria from a list of 20 FOSS ILS. Ahammad (2014) describes implementation of Koha ILS with all its modules at the Independent University, Bangladesh.

The present study is based on the experiment carried out by authors at Tezpur University library. The main objective of the study is to extract various kinds of data from existing ILMS Libsys, transform those data as desired by new ILMS Koha, migrate transformed data to new ILMS Koha, and validate migrated data of Koha with Libsys. This paper and the findings of the study will encourage not only librarians but also those who are involved in library automation project, particularly data migration from one ILMS to another ILMS.

4. Methodology

The entire process of data extraction, data transformation and data migration was accomplished by using LIBSYS, MS-Excel, MarcEdit and the Koha ILS. The mechanism for data migration from Libsys to Koha is depicted in Figure 1.

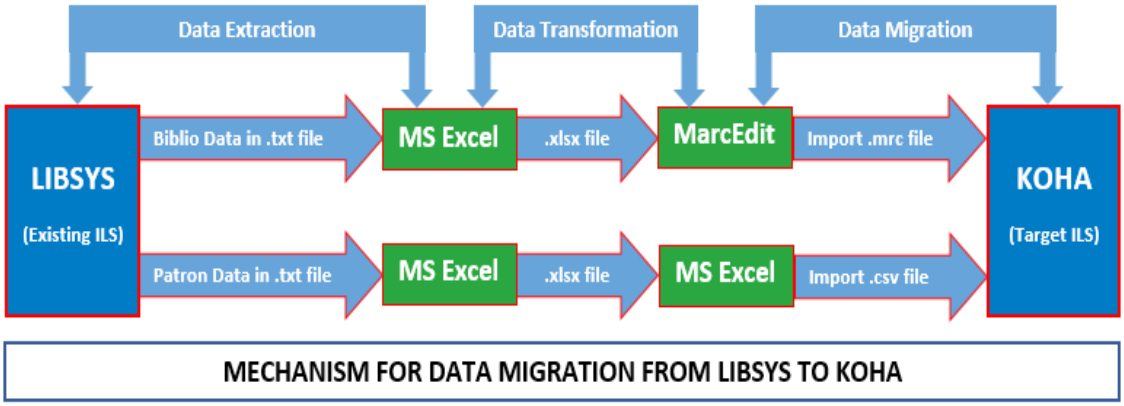

Figure 1. Mechanism for Data Migration from Libsys to Koha.

4.1 Data Migration Process

The procedure for data migration from one system to another system was accomplished in four different stages such as data extraction, data transformation, data migration and data validation. Before doing data extraction, we examined both the library management systems i.e. LIBSYS and KOHA. We observed that the target ILMS Koha desires bibliographical data in MARC (.mrc) format and patron data in CSV (.csv) format, but the existing ILMS Libsys provides both bibliographical data and patron data in text (.txt) format. Hence, the text format of both data needs some transformation. After transformation of text data format to the new data format desired by the target ILMS, we initiated migration and validation of data.

The four stages of data migration procedure are elaborated:

4.1.1 Data Extraction

Data extraction is the process of extracting various kinds of data i.e. bibliographical data and patron data from an existing ILMS (Libsys). For this purpose, two checklists were prepared, viz., (i) List of the data fields for the bibliographical record, and (ii) List of data fields for the patron record as per the requirement of target ILMS Koha. Then all options and all modules

of Libsys were explored to get maximum possible data of each list. In Libsys, export option in the cataloging module does provide sufficient bibliographical data; but the vendor disables it before installing Libsys. Hence, data extraction using export option in Libsys to get all bibliographical data was not possible. By studying some research and by a system analysis of Libsys, it was found that accession register in the acquisition module of Libsys provides maximum possible bibliographical data and reports on members in the circulation module of Libsys provides maximum possible patron data. By using the Checklist-I and Checklist-II, we extracted all bibliographical data and patron data in text format.

Check List I. List of data field for the bibliographical record

Sl. No.	Data Field Name	Description
1	ISBN/ISSN	International Standard Book/Serial Number
2	Call Number	Dewey Decimal Classification No. + Book No.
3	Main Entry-Author name	First Author or Corporate Author as Main Entry
4	Title	Title of the Document
5	Statement of Responsibility	Statement of Responsibility of Document
6	Edition	Edition of Document
7	Imprint	Place and Name of publisher, Year of Publication
8	Physical Description	No. of pages, illustration, size, etc.
9	Series	Series Statement
10	Notes	General Note
11	Subject Heading	Subject Heading
12	Added Entry-Author name	Author, Editor, etc. Added Entry
13	Document Type	Type of document available
14	Shelving Location Code	Document type shelving location code
15	Date of acquisition	Date of acquisition of document
16	Source of acquisition	Coded value of the Vendor
17	Purchase price	Purchase price or List price of document
18	Accession number	Accession Number

Check List II. List of data field for the patron record

Sl. No.	Patron Data
1	Patron Category
2	Library Card Number
3	Full Name of Patron
4	Date of Birth
5	Gender
6	Department Address
7	Alternative Address
8	Email
9	Mobile
10	Enrollment Date
11	Expiry Date

4.1.2 Data Transformation

Data transformation is a process of transforming data from one format to another format as desired by the new system. In the data extraction stage, individual text files of bibliographical data and patron were collected. These data cannot be directly accepted by the target ILS Koha. Hence, these needed to be transformed into the desired format. In this stageo begin with, both the text files were converted into excel file format. All the bibliographical data was then thoroughly examined to remove all redundant and invalid data. Additional bibliographic data as required by the target ILMS was added and the data rectified as per the requirement of target ILS Koha. The corresponding MARC tags of Koha were added in the first row of excel file (e.g. 020$a for ISBN, 080$a for DDC no., 100$a for main entry-author name, etc.) and finally, the excel file was imported into MarcEdit software which resulted in a .mrk file which was then compiled to .mrc file. This .mrc file containing all bibliographical data is ready for import into the target ILS Koha.

Similarly, patron data was examined and all inconsistent data was removed, required additional data was added and data was rectified as per the requirement of target ILS Koha. This excel file was converted into .csv file. This file is now ready for import into the target ILS Koha. Figure 2 depicts data transformation process from text file format to MARC file format.

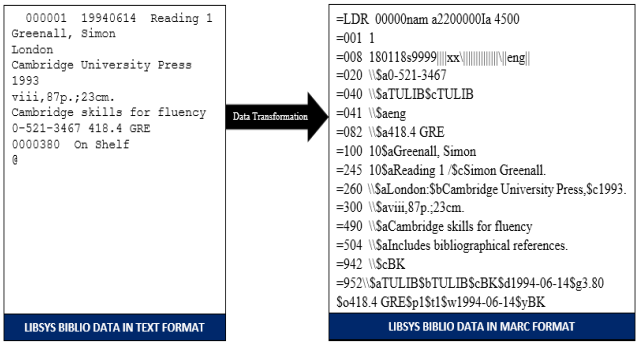

Figure 2. Data Transfer: Text to MARC format.

4.1.3 Data Migration

Data migration is the process of transferring various kinds of data from the source system to the target system. Before starting data migration, Koha 18.11 on Ubuntu 16.04 server as per the requirement of the library were installed and customized. In the new ILMS Koha, a library with different item types and their shelving locations, patron categories and attribute types, circulation and fine rules were created, and other parameters and administrative system preferences, etc. as per the Koha implementation checklist available in Koha manual were set. Once all modules in global system preferences of Koha ILS are set, then the process of data migration starts. Koha server is now ready for import of both bibliographical data and patron data.

At the end of data transformation stage, all bibliographical data in MARC and all patron data were transformed into .mrc format and comma separated value (.csv) format respectively. Both these data formats were now ready to be imported to the target ILS Koha separately using the tools module of Koha. There are two ways to import bibliographical data (.mrc file) to Koha ILS server. One is through Graphical User Interface (GUI) of Koha ILS server and another one is through command line interface (CLI) of Koha ILS server. Hence, transferring bibliographical data (.mrc file) and patron data (.csv file) through GUI of Koha ILS is easy but time consuming. In this project, for data migration, CLI was carefully used to import all bibliographical data and GUI was used for import of all patron data to new ILMS Koha.

4.1.4 Data Validation

Data validation is the process of verifying imported bibliographical data and patron data in the new system (Koha) and checking and comparing it with the data in the existing system (Libsys) for accuracy. This process is the final stage of data migration process. In this study the bibliographical data and patron data was validated by randomly using staff client interface and WEBOPAC interface of both Libsys and Koha.

4.2 Comparison of Existing ILMS and Target ILMS

Characteristics	Existing ILMS	Target ILMS
Name of ILMS	Libsys	Koha
Version	4 (Old version)	18.11 (Latest stable version)
Type of Software	Proprietary	Open Source
Application Type	Desktop based application	Web based application
Ownership	Libsys Ltd., Gurgaon, India	Katipo communications Ltd., Wellington, New Zealand
License	Commercial	Under GPL General Public License
Price	In Lacs	Freely available and free support
Customization	Libsys charge users to provide customized solutions	Koha provides source code, which is freely available for innovation to provide new features at users end.
Training manual	No system manual is provided to users except user manual to get AMC	Koha manual includes everything for user convenience

Web Server	Only Apache and IIS	Apache, IIS and others
Database	Own database used for storing all kinds of data. But Libsys7/Librsys10 Software used SQL Server/ORACLE/ MySQL as a backend RDBMS	Uses MySQL/MariaDB RDBMS. Scalable enough to meet the transaction load of library.
Vendor Lock-in	Restrictions – can ask for support only from particular vendor	No restrictions, no set term contracts on changing support
Data in MARC format	Don't provide data in MARC format	Provides data in MARC format
Unique features	Don't support offline circulation	Supports Offline circulation
Reports	Limited report library	Large report library
Addition of new features/new version	Charge extra cost to upgrade to new version or add new features	New versions are coming in every six months and added for free

5. Findings

The document types available in the existing system Libsys are textbooks, reference books, gifted books, law books, thesis, dissertations, CD/DVDs, conference proceedings, government publication books, etc. The same document types were created in the new system Koha. Table 1 shows that the number of textbooks + reference books, gifted books, law books, thesis, dissertations, CD/DVDs/ VHS Cassettes, conference proceedings, government publication books extracted from Libsys and imported to Koha are 78611, 3834, 297, 507, 1300, 2287, 75 and 569 respectively. During data validation, it was found that the numbers matched clearly indicating that all 87480 bibliographical records have been successfully migrated to the new ILS Koha. Data validation process was accomplished by using Koha staff client interface and Koha WEBOPAC interface.

Table 1. Document types

S. No.	Document Types	No. of Biblio records extracted from Libsys	No. of Biblio records migrated to Koha	Data Validation
1	Text & Reference Books	78611	78611	Successful
2	Gifted Books	3834	3834	Successful
3	Law Books	297	297	Successful
4	Thesis	507	507	Successful
5	Dissertations	1300	1300	
6	CD/DVDs/VHS Cassettes	2287	2287	Successful
7	Conference Proceedings	75	75	Successful
8	Govt. Publication Books	569	569	Successful
Total		87480	87480	Successful

Similarly, the patron categories available in the existing system Libsys are student, research scholar, teacher, staff and project fellow. The same patron categories were created in the new system Koha. Table 2 shows that the number patrons of the library such as student, research scholar, teacher, staff and project fellow extracted from Libsys and imported to Koha are 2806, 550, 283, 228 and 6 respectively. During data validation, we found that the same number of individual patrons were available in Koha. Patron data validation accomplished by using Koha staff client interface indicated that the data validation process of all patrons was successful.

Table 2. Types of Patrons

Sl. No.	Patron Category	No. of Patron records extracted from Libsys	No. of Patron records migrated to Koha	Data Validation
1	Student	2806	2806	Successful
2	Research Scholar	550	550	Successful
3	Teacher	285	285	Successful
4	Staff	228	228	Successful
5	Project Fello w	6	6	Successful
Total		3875	3875	Successful

6. Conclusion

The migration of data from one software to another software is always a tedious job. The study found that both bibliographical data and patron data in the new system (Koha) are identical with what is available in the source system.

7. References

Avery, J. M. (2016). Implementing an open source Integrated Library System (ILS) in a special focus institution. *Digital Library Perspectives*, *32*(4), 287–98. https://doi.org/10.1108/DLP-02-2016-0003

House, M. D. (2016). Implementing the open-source Koha-ILS at the Deutsche Schule Charlotte. *Digital Library Perspectives, 32*(4), 253–69. https://doi.org/10.1108/DLP-02-2016-0007

http://epgp.inflibnet.ac.in/epgpdata/uploads/epgp_content/library_and_information_science/information_communication_technology_for_libraries/25._case_study_libsys_/et/2089_et_et.pdf

http://katipo.co.nz/

http://manual.koha-community.org/

http://schema.koha-community.org

http://www.libsys.co.in/

http://www.loc.gov/marc/bibliographic/

https://blog.reeset.net/archives/2686

https://en.wikipedia.org/wiki/MarcEdit

https://en.wikipedia.org/wiki/Microsoft_Excel

https://koha-community.org/manual/17.11/html/15_implementation_checklist.html

Müller, T. (2011). How to choose a free and open source integrated library system. *OCLC Systems and Services: International Digital Library Perspectives, 27*(1), 57-78. https://doi.org/10.1108/10650751111106573

Pund, M., and Jain, P. (2016). System and process for data transformation and migration from Libsys to Koha. *International Research Journal of Engineering and Technology, 3(4)*, 690–702.

Todd, C. R. (2018). Librarian as data migrator: A functional pathway from Millennium to Koha. *Digital Library Perspectives, 34*(1), 60–9. https://doi.org/10.1108/DLP-09-2017-0035

Youn, C., and Ku, C. (1992). Data migration. *Bell Communications Research, Piscataway*, 1255–8.