



Data Migration from Libsys to Koha at State Central Library, Kerala: A Preliminary Assessment

Jamsheer N P, Manu C and Mohanan A

Abstract

Migration of Library Management System (LMS) without the loss of valuable data is a tedious task. There is challenge in data migration, but it also offers opportunity for rectifying existing problems and redesigning the workflow of the library. During the process of system migration a number of improvements and catalog enrichments can be done. Typographical errors and other problems of the catalogue can be rectified during the process of data migration. The article discusses the problems in library catalogue, challenges and opportunities in data migration and various steps to be taken during the system migration. It deals with the feasibility of migration from LibSys to Koha based on the existing system in Kerala State Central Library, Trivandrum one of the largest collections in the country. The study suggests various check points to be fulfilled during the process of migration. The authors themselves carried out the system evaluation and feasibility analysis for the migration project and have first hand information regarding the process.

Keywords: Library, Automation, software, LMS, ILMS, Koha, LibSys, Data export, MARC21, Metadata Standards.

Introduction

Availability of free and open source software like Koha have provided libraries around the world with the confidence to migrate from existing proprietary Library Management Systems (LMS) to free software. A Library Management System accumulates various types of strategic data sets over the years. The first step of successful migration of library system is the identification of these data elements, documentation and designing of a detailed workflow for exporting these legacy data. Complete data extraction is probably the most important and certainly the difficult aspect of the whole system migration process. It is also the sensitive part

as mistakes made in data migration can be profound and unrecoverable (Seaman 1996).

Various libraries including major university libraries and public libraries in India have successfully migrated from various proprietary software to Koha. According to Pourciau(1992) a library's decision to migrate from a legacy system to a new integrated library management system depends upon various factors such as lack of functionality, hardware compatibility issues, unresolved system problems etc. Financial burden of managing a proprietary software is one of the main factor for migrating from proprietary to open software in various libraries.

Migration needs to be done without loss of any valuable data accumulated over years. Therefore, objectives of a migration programme can be listed as follows:

- Avoiding unacceptable risks associated with using of a near obsolete version of LMS to support critical library automation activities and migrate to a new system.
- Acquiring additional features and to continue to maintain the productivity levels that were achieved by the installation of the existing Integrated Library System.
- Preserving and building upon the existing bibliographic records and to make them compliant to international bibliographic standards.

Koha: Development Perspective

Koha is an open source Integrated Library System (ILS), used worldwide by public, school and special libraries. Koha is a true enterprise-class ILS with comprehensive functionality including basic and advanced options. Koha includes modules for acquisitions, circulation, cataloging, serials management, authorities, flexible reporting, label printing, multi-format notices, offline circulation for when Internet access is not available, and much more. Koha will work for consortia of all sizes, multi-branch, and single-branch libraries.

Koha has facility for using numerous languages, and new ones are added every year which makes it Multilingual and translatable. It has a powerful search interface, and an enhanced catalogue display that can use content from Amazon, Google, LibraryThing, Open Library, and Syndetics, among others. Koha's OPAC, circulation management and self-checkout interfaces are all based on standards-compliant to World Wide Web technologies—XHTML, CSS and Javascript – making Koha a truly platform-independent solution. Koha is built using library standards and protocols such as

MARC 21, UNIMARC, z39.50, SRU/SW, SIP2, SIP/NCIP, ensuring interoperability between Koha and other systems and technologies, while supporting existing workflows and tools.

Koha is distributed under the Free Software General Public License (GPL) version 3 and hence qualify as a true free and open source software (FOSS). Libraries are free to install and use Koha themselves if they have the in-house expertise or to purchase support or development services from the best available sources. The availability of community support as well as commercial support eliminates problems such as vendor-lock in.

Koha is web-based Integrated Library System (ILS), with a SQL database backend and cataloguing data stored in MARC and accessible via Z39.50 or SRU. The user interface is customizable and can be adopted for requirements of various types of organizations. As it is Unicode compatible it can be used in most of the language environments of the world. Koha has most of the features that would be expected in an ILS, which include the following:

Free software like Koha helps librarians serve the public on a reasonable budget. Proprietary ILS packages are very expensive for purchase and maintenance as well as updation. A larger library like Kerala State Central library with good collection has to bear huge annual maintenance cost, as many charges are based on quantum of collection. Once a library has bought a proprietary system, they experience a high barrier to change as well. Data is most often kept in proprietary formats from which it is difficult to export - in some cases, the data is actually "owned" by the system vendor.

As with all non-free software, customers are left at the mercy of their vendors for enhancements and customizations. Library system vendors historically have been slow to provide innovative new options.

Although user groups exist for many of the existing systems, they seem to be more like mutual support groups than sources of feedback for the vendors. A worse fate is in store for those whose ILS vendor goes out of business or is bought by another vendor. This situation presents a great opportunity for free software. It's an opportunity that's not lost on librarians either, at least not all of them. There is still a great deal of ignorance and inertia to overcome.

A lot of opportunities are on the horizon for Koha in India. Government of Kerala in India has accepted and approved it giving a status very much like an official software. Kerala State Library Council has accepted Koha as standard and about 8000 rural libraries existing in Kerala will be automated using Koha customised for the regional language; Malayalam.

Already Koha is installed in many university libraries of Kerala with each having collections of more than 1,00,000 items. The search tools are getting customized to improve their efficiency and allow searching options for regional language Malayalam.

Koha is currently a very active project. It has a very large, active development team and a mature, well-established codebase. The latest stable release of Koha is expected at the beginning of 2017.

Kerala State Central Library

Trivandrum Public Library was the premiere public library in Kerala, established in the year 1829, during the reign of King Swathi Thirunal of Travancore. The Trivandrum Public Library was declared as State Central Library (SCL) in the year 1958. The SCL has received the entire collection of most of the Dewans and other higher officials and many scholars from Kerala and recently the collection of British Library, Trivandrum (2005) on its closure. The SCL has total collection of 3,60,000 books and 275 journals and functions in the heritage building.

Children's library functions in a new building designed very similar to the earlier heritage structure.

The library uses the subject classification scheme Colon Classification (CC) 6th edition for the main library as well as children's library and Dewey Decimal Classification (DDC) scheme for the books in the British Library Collection. The existing automation system software has not fully incorporated MARC21 cataloging standard.

SCL has been using Libsys software as library management system from the year 2005 running on Microsoft Windows server platform. Libsys is a widely used proprietary software developed in early 90s by Libsys Ltd, New Delhi, India and has all the modules required to manage libraries. The SCL has integrated RFID for circulation of books and serials. Now, SCL has decided to migrate to Koha; a free and open source LMS. Koha is a fully featured, scalable LMS developed by HLT, New Zealand and supported by developers all over the world.

The functional units of the library and their corresponding implementation in automation system are analysed in the following sections which are considered for formulation of a proper migration plan.

Bibliographic Records and Holding Information

Catalog is the core product of a library and migration should be done without loss of any of these data accumulated over the years. Libsys version 4 used in SCL does not provide any comprehensive export options to extract complete bibliographic data in standard formats such as MARC21 or CSV (comma separated value). Analysis of the existing catalogue has revealed that thirty three data fields are to be identified as important and exported as such in order to migrate data successfully. The fields that need proper extraction are given in Table 1 and table 2.

Table 1. Minimum bibliographic data fields to be exported from Libsys

<ul style="list-style-type: none"> • Class Number • ISBN/ISSN • Main entry - Personal name • Main entry - Corporate name • Added entry - Personal name • Added entry - Corporate name • Meeting • Title • Remainder of title / Subtitle • Edition • Statement of responsibility • Place of publication 	<ul style="list-style-type: none"> • Name of publisher • Date of copyright • Extent (pagination) • Other physical details • Dimensions • Series statement • General Note • Language note • Subject / topic / keyword • Abstract • Volume Information
--	---

Table 2. Details of Holding / Item Data Fields to be Exported

<ul style="list-style-type: none"> • Accession number • Circulation status • Type of document • Full call number • Collection details 	<ul style="list-style-type: none"> • Location information • Item Price • Withdrawal status • Binding information • Any other local information
--	---

Membership and Circulation

SCL has more than 86000 library members, under five major categories. These membership details are another important data set to be migrated to the new system. Membership details stored in 14 different data fields were identified for migration (see table Table 3). In addition to the bibliographic information and patron data, the circulation data also needs to be migrated. Library circulates an average of 2500 books and serials per day as issue, return and renewal. Details of books and serials issued to various members, date of issue and due

date are required in the circulation module. Library automation system has generated piles of circulation history over the time. This valuable information can be used later for streamlining acquisition policies and for user studies. Libsys has the option to export circulation history and Koha allows migration of old circulation data if necessary.

Analysis of Data in the Current System

The quality of data in the existing system needs to be analyzed before the actual migration. A clear understanding of the existing data can save time and energy during

Table 3. Membership fields to be exported from Libsys

<ul style="list-style-type: none"> • Patron ID (Card number) • Patron Name • Patron Category • Gender • Date of birth • Full Address • Full Address - second 	<ul style="list-style-type: none"> • Email • Phone number — Residence • Phone number — Office • Registration date • Date of membership expiry • Cancellation Status • Name of nominee
---	--

the process of migration. Cleansing of messy data require a lot of manual as well as semi automated labour, as fully automated tools cannot completely replace the knowledge and experience of librarians. Typographical errors and non standard rendering have always been a problem in the LMSs. Various studies have tried to analyse the problems related with typographical errors in bibliographic records. Beal and Kafadar found erroneous bibliographic records can cause retrieval problems in online catalogue (Beal and Kafadar, 2004). According to Anne Christensen (2013) early online catalogues served internal purposes in the first place, rather than being meant as a service for users (Anne Christensen, 2013). During the initial stages of library automation, data entry work was done by non professional staff. In the era of web OPAC and discovery tools, these erroneous data would compromise precision of the search. Some of the examples of commonly repeated erroneous and non-standard data in library catalog are given in Table 4.

The existing non standard data like the above can be standardized and typographical errors

can be corrected during migration process. Catalogue can be enriched by remapping data entered in different fields and by adding additional information. Existing name of authors, publication details etc. rendered differently (see table 4) can be standardized. To avoid such problems in the future, library can make use of a name authority file.

Enrichment and Merging of Catalog.

Koha follows MARC21 as its cataloging standard. In addition to the common bibliographic fields, MARC21 contains various fixed length fields for recording additional information such as form, genre and language information. The existing catalog of SCL has not included these types of information. These can be included during the post processing for migration. If enhanced, the search will retrieve and display relevant records in the online public access catalog with all the required pieces of information.

SCL houses a large collection of titles with multiple copies. Most of these books are cataloged as multiple holding of the same bibliographic record but there are large

Table 4. Example of data rendered differently in various fields

Author	Place of Publication	Publisher	Title
Kurup, O N V	New Delhi	Rupa &Co	Nilavilikalkumel Aarute Nissabdatha—?(M)
Kurup O N V	New delhi	Rupa Co	Nilavilikalkumel Aarute Nissabdatha?
O N V Kurup	New Delhi	Rupa &Co.	(M)Nilavilikalkumel aarute nissabdatha (M)
KURUP, O N V	New DELhi	Rupa Co.	Nilavilikalkumel aarute nissabdatha—? (M)
Kurup ,O N V	new Delhi	Rupa &co	Nilavilikalkumel aarute nissabdatha? (M)
Kurup, O. N. V.	new delhi	Rupa &Co	
kurup, O. N. V.	NEw Delhi	Rupa. Co	India: 50 years of Independence; 1947-97
	NeW Delhi	Rupa &Co.	India: 50 years of Independence: 1947-97
	N Delhi	Rupa &co.	India; 50 years of Independence: 1947-97
	N delhi	Rupa .Co	India 50 Years of Independence 1947-97
	N DELhi	Rupa&Co	India: 50 years of independence 1947-97
	N. Delhi	Rupa &CO	India: 50 years of independence; 1947-97
	new Dalhi	Rupa &CO.	India; 50 years of Independence; 1947-97
	naw delhi		India; 50 years of independence 1947-97

numbers of duplicate entries too in Libsys database. Duplicate bibliographic records should be converted into single bibliographic record by merging for easy management of the catalog and for better search precision. Koha has a provision for manually merging these records one by one, but it would take much time especially in libraries having large collection. The benefits of bibliographic record merging can be as follows:

- Easy cataloging as existing bibliographic record need not be recreated.
- Improved search precision.
- Accurate reporting.
- Better performance in Koha's inbuilt "Refine your search" option.
- Helps formulate acquisition policies
- Easy book reservation and retrieval

Pre-requisites

During migration to an integrated library system from LibSys to Koha, SCL need to pay attention on the following:

- Provide detailed in-house Koha training to the library staff before the actual migration.
- Conduct rigorous test in Koha with sample data exported from Libsys to make sure that all modules are working properly.
- Decide the bibliographic standards to follow in cataloging.
- Check the hardware and software compatibility of existing RFID devices with new system.
- Since post migrated environment may bring new challenges due to changes in the workflow, sufficient technical support may be ensured.
- Explore and utilize the new features available in Koha like bulk import of patron images, Email notification, SMS notification etc.
- Develop responsive web based OPAC for both computers and mobile devices.

9. Conclusion

Academic and public libraries including major university libraries are migrating to Koha all over the world. It helps libraries to make resources discoverable and accessible to the wider user community and to cope with rapidly changing technology and increasingly sophisticated users. The libraries intending to shift to Koha needs to fix prerequisites of migration processes and should develop and pursue a proper action plan as discussed above. Normalization of differently rendered entries and merging of similar titles would provide efficient library management and better user experience. The library can also save money, time and achieve efficiency in services.

Reference

- Beall, J, & Kafadar, K (2004). The effectiveness of copy cataloging at eliminating typographical errors in shared bibliographic records. *Library Resources & Technical Services*, 48(2), 92-101.
- Christensen, A. (2013). Next-generation catalogues: what do users think. *Catalogue 2.0: the future of the library catalogue*, 1-15.
- Eyler, Pat (2003). Koha: a Gift to Libraries from New Zealand. International. Feb 01, 2003 Koha community web page (<https://koha-community.org/about>)
- Pourciau, L J (1992). Automated library system migration in the United States. *The Electronic Library*, 10(2), 103-108.
- Ransom, Joann (2009). HYPERLINK "<http://library-matters.blogspot.com/2009/09/liblime-forks-koha.html>" Library Matters: Liblime forks Koha HYPERLINK "<http://library-matters.blogspot.com/2009/09/liblime-forks-koha.html>". library-matters.blogspot.com. Retrieved 2015-08-10.
- Seaman, Scott (1996). Circulation and data migration: a CARL systems to innovative interfaces case study. *Library and information Science Research e-journal*, 6(2).
- State Central Library website <http://statelibrary.kerala.gov.in/history.htm> (retrieved on 06th December 2016)
- Wikipedia (2016). [https://en.wikipedia.org/wiki/Koha_\(software\)#cite_note-36](https://en.wikipedia.org/wiki/Koha_(software)#cite_note-36). Accessed on 03.12.2017