

4-29-2019

Data Migration From Legacy Systems To Koha: A Practical Approach

Vimal Kumar

Mahatma Gandhi University, vimalibre@gmail.com

Majeed Abdul

Farook College, kcamajeedfc@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>

Part of the [Library and Information Science Commons](#)

Kumar, Vimal and Abdul, Majeed, "Data Migration From Legacy Systems To Koha: A Practical Approach" (2019). *Library Philosophy and Practice (e-journal)*. 2559.

<https://digitalcommons.unl.edu/libphilprac/2559>

Data Migration From Legacy Systems To Koha: a practical approach.

Vimal Kumar V.

Research Scholar

School of Social Sciences

Mahatma Gandhi University,

Kerala-686 560, India

Email: 2808@mgu.ac.in

Dr. K. C. Abdul Majeed

Associate Professor

Farook College, Kozhikode-673 632

Email: kcamajeedfc@gmail.com

ABSTRACT

Koha library management software opened up a new ray of hope for libraries not automated yet in India. Many libraries would like to free from proprietary software and switch over to Koha. Data migration from legacy software, not an easy task due to vendor lock-in and data in a non-standard format. Good homework and teamwork required for the successful migration from proprietary software to Koha. The article gives an overview of various processes in data migration from legacy library automation software to Koha.

Keywords: Data migration; Koha; Open Source Software; Integrated Library Management System; Library automation.

INTRODUCTION

The dominance of proprietary software in the market is one of the main reasons for libraries not automated in developing countries. High price, lack of control over data, stringent license agreements, and proprietary supporting software (e.g., database, operating system) result in increased total cost ownership of library automation software. Libraries with essential IT infrastructure could automate the library services with the help of Open Source Software like Koha. Koha software shows maturity in

features and functions and suitable for any types of libraries. Libraries with no automation software can directly implement Koha and start the automation processes from scratch. Libraries using proprietary software and wish to move to Koha in a dilemma due to data export from vendor locked legacy system. Legacy automation systems with unstructured data also create problems during data export.

Earlier generations of library management software developed to address the requirements of the printed collection. Nowadays integrated library management software need to represent both print and digital resources. Integrated library systems should address the requirements of cooperation and resource sharing among libraries. Libraries expect the features like cooperative collection development, discovery services, interlibrary loan, collection analysis, sharing of print and digital resources with library management systems to meet the user requirements (Machovec 2014)². Modern integrated library systems characters are cloud-based, security tightened, social media integrated with mobile and web interfaces. Koha almost fulfills the requirements of libraries in a networked environment. Koha has the provision and capability to work with add-on modules to extent the services.

Many libraries with national importance could migrate to Koha in India. They could migrate to Koha with the help of either using in-house expertise or hired expertise. Such instances have increased the popularity of among libraries in India. Libraries using proprietary solutions started to implement Koha. However, they have limitations in move data from legacy systems due to various reasons like vendor lock-in, lack of technical expertise, and data in an unstructured/strange format. Data migration from proprietary library automation systems is a challenging task. Library professionals need to familiar with the data structure, mapping and various phases in the migration processes. Then only library professionals can lead the migration process to Koha successfully without any data loss.

REVIEW OF LITERATURE

The literature review on various aspects of library automation and Open Source library management systems indicate that Koha is the popular library automation software among libraries. Migration from legacy automation system to Open Source one is a challenging task.

Bhardwaj and Shukla discuss certain factors which force library automation. The factors are Information explosion, shrinking library space, the necessity to fulfill the new desires of users, sharing of resources, and assistance for library automation (Bhardwaj and RK Shukla 2000)^[2]. Anil Singh states that computerization of library services in India was very slow and only 3% of libraries could automate at the end of 1990's (Singh 2003)^[2]. With the assistance of INFLIBNET Center, 142 universities could develop infrastructure and received training for library automation in the financial year 2000-2001. The SOUL software version 1 released in 2000. Many academic libraries could automate the library services with the help of SOUL software at a minimal charge (Chandrakar and Arora 2009)^[2]. Recent studies on library automation indicate that Koha is popular among all types of libraries in India. The number of libraries automated using Koha has increased (V. Kumar and Jasimudeen 2012)^[2](Khode and Chandel 2015)^[2](T. G. Kumar and Jayapradeep 2015)^[2]. A study conducted in 2003 to find the constraints in library automation among library professionals at central universities. The study has listed out the constraints in the Indian context. The highly ranked constraints are inadequate financial resources, lack of well-accepted standard software package and non-availability of IT trained personnel (Ramana, Information, and 2003 n.d.)^[2].

Library automation using Open Source Software can provide efficiency and effectiveness at a minimal cost (Ukachi, Nwachukwu, and Onuoha 2014)^[2]. Open Source library management systems considered as an alternative to proprietary software in libraries. Development and maintenance costs of Open Source

alternatives lower than commercial alternatives. Koha is the popular Open Source library management system with all functional modules. Koha supports the regional languages in India due to the availability of Unicode. Saxena and Srivastava share the result of an evaluation of library software popular in India in 1998. At that time proprietary software ruled the library automation market. Such systems require other proprietary components like operating system, database management system, and antivirus software (Saxena and RK Srivastava 1998) [2].

Balnaves reports that Open Source library management systems show steady improvements in features and functions. The Open Source software projects focus on community participation, open source code, and database schema and dimension (Balnaves 2008) [2]. Singh and Sanaman attempt to evaluate an earlier version of Koha and point out the advantages. Koha meets all specifications of Open Source Software than other counterparts. The documentation of Koha is very comprehensive. Koha supports more standards and formats match with industry standards.

Various costs involved in the migration of legacy library automation to the new one. Following costs involved in the migration process; workforce, conversion of bibliographic, user and transaction details (Sahu, Hemant Kumar, N. Nageswaran 2005) [2]. Karak and Dutta share the experience to convert bibliographic details from spreadsheet to Koha. Libraries without any automation software keep their data in the spreadsheet. Such libraries can quickly move data to Koha Open Source library management software (Karak and Dutta 2017) [2]. Chattopadhyay and Sarkar give a detailed account of data migration from LibSys library management system to Koha. They exported bibliographic details from LibSys in text format. Then converted the bibliographic details suitable to Koha with the help of MarcEdit software. Proprietary software does not provide any option for data export and migration become a challenging task (Chattopadhyay and Sarkar 2017) [2].

The review of the literature gives an overview of library automation scenario in libraries. Proprietary systems dominate the global library automation market. Such systems are difficult to maintain because of expensive to implement and maintain. Open Source library management systems have been creating a presence in all types of libraries. The article on guidelines on migration to Open Source software can enrich the literature in this domain and helpful for many libraries.

KOHA AND LIBRARIES IN INDIA

Academic and research libraries in India run by public fund enjoy the benefits of automation using integrated library management system. In certain extent, university and college libraries could automate with the support of SOUL software developed by INFLIBNET Center, Gandhinagar, Gujarat. Many libraries with the small collection have selected e-Granthalaya software of National Informatics Centre. A good number of other libraries had automated using library automation software developed by local vendors. A minority of libraries with proper financial assistance could afford comprehensive library automation solutions. Libraries not able to afford a library management software could manage bibliographic details with Information Storage and Retrieval systems like CDS/ISIS. CDS/ISIS was an excellent choice for libraries with minimal IT infrastructure to start the library automation. However, libraries can export the bibliographic details from CDS/ISIS to integrated library management systems without any restrictions. Majority of Indian public libraries are not in a position to enjoy the benefits of automated library services due to lack of professionalism and infrastructure.

Following are the key reasons to switch over from legacy automation software to Free and Open Source Software in the context of developing countries like India:

- ❖ Shortage of fund to maintain the proprietary software service.

- ❖ The inability of the legacy system to meet the expectations of the library community.
- ❖ Lack of control over the software and library data.
- ❖ Lack of regular software updates.
- ❖ Obsoleted technology platform of legacy automation system.

Koha noticed among Indian libraries after the adoption at Delhi Public Library in 2008. The collection strength of the Delhi Public Library is over 1.5 million records. The library could port a small portion of the bibliographic details stored in CDS/ISIS to Koha in the initial stages of migration (Breeding 2008)². Many libraries with funding and inhouse expertise switched over to Koha in India. Libraries of IIM, IIT, university and research libraries in India had selected Koha in the early years of adoption. A significant number of academic libraries in the Kerala state migrated to Koha by utilising the expertise of the Free Software Community. Koha became popular among libraries in West Bengal with the support and training from Bengal Library Association.

THE MIGRATION PROCESS

The transition to Koha is comparatively easy for libraries with technical support. Koha offers various conveniences for those who want data migration. Availability of various data import options (GUI and command line tools), compatibility with bibliographic standards, open database schema, documentation, community support make data migration easy to Koha. Expertise in Linux based operating systems, database management, familiarity with library standards and workflow are the skillsets required for the person who handles the data migration. Following are the various phases of data migration from legacy automation system to Koha:

- ❖ The organisation of a feasibility study. It includes the checking of the suitability of the proposed system in the organisation.
- ❖ Permission from authority and seeking options for financial assistance.

Communication with the legacy software vendor regarding stopping of service and data export.

- ❖ Arrangements for the documentation of the migration process.
- ❖ Purchasing and preparing infrastructure including computer hardware, hosting space, and networking facilities for the new system.
- ❖ Preparation for data export from the legacy automation system.
- ❖ Creating a testing environment using Koha with sample data.
- ❖ Training for library staff.
- ❖ Importing of data for the final stage.
- ❖ Arrangements for data security including backup, implementation of network security etc.
- ❖ Hosting of Koha and make available OPAC for the public.

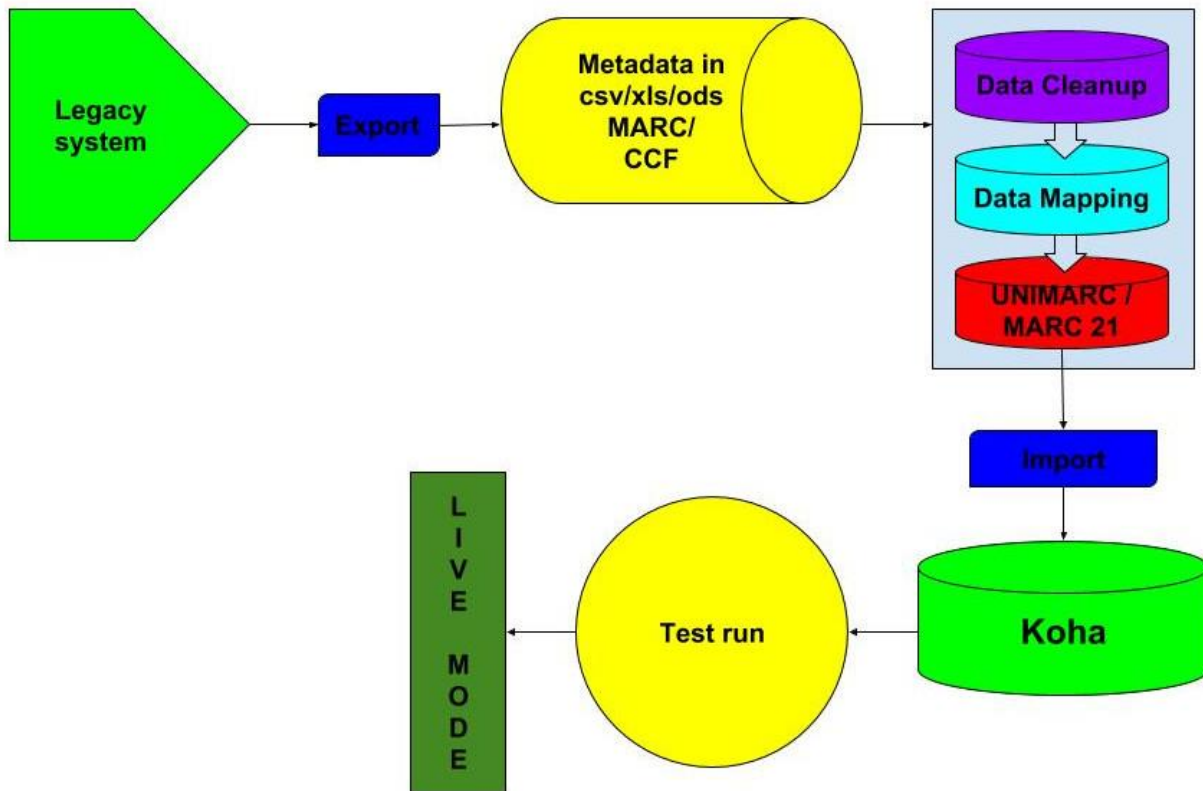


Fig. 1. Data migration processes to Koha

Following are the detailed account of the data migration process to Koha:

Planning stage

A library wishes to implement Free, and Open Source Software has to consider the availability of in-house technical support. Libraries are traditionally familiar with paid support from the software vendor. Library professionals need to familiar with new Open Source Software tools to manage various occasions. In certain instances, libraries need to invest more in hardware and to seek external support to implement and maintain the new system.

First, assess the library needs, evaluate the FOSS software candidate and then decide whether it right for the library. Find the strength and weakness of the library regarding budget, workforce, and IT infrastructure. Following are the various elements to be considered in the planning stage of the Free and

Open Source library management system:

- ❖ Permission from authority to implement a new library management system.
- ❖ Create consensus among library staff on the replacement of library automation software.
- ❖ Assessment of infrastructure required for the implementation of the new library management system.
- ❖ Decisions regarding type of software service support; vendor or community support.
- ❖ Planning on training and manpower development.

Libraries in developing countries face more or less similar forces which slow down the automation processes. Certain factors which affect the library automation emerge from inside the organization itself. They are lack of willingness, unity, and difference of library staff. Inadequate funding, lack of coordination, delays due to red rapism, non-supportive attitude from decision makers and

administrative staff also slow down the implementation of cutting-edge technologies [Balaji]. Various measures need in advance to avoid the disturbances which may affect the routine library activities during the migration period. A testing environment should create to try exported data with the proposed software. Koha can run parallel with legacy system and give the opportunity for library staff try all functional modules. Arrange occasions for library staff to visit the libraries successfully using Koha and interact with the users there.

Data export

Exporting of bibliographic, patron and circulation data from legacy systems is a challenging task. Very few library management systems follow standards for bibliographic details like MARC and CCF. The success of data export depends on whether the software vendor allows exporting of data. The previous version of LibSys library software had the provision to export the data to MARC format. No provision to export the bibliographic data from the latest versions of LibSys. SOUL software allows exporting of bibliographic details to MARC and CCF format. Very few proprietary software vendors have made available options in software GUI to export bibliographic data.

The exported bibliographic details in a standardized format like MARC or CCF reduce the efforts of data migration. Systems like SOUL version 1 allows exporting of bibliographic details to either MARC 21 or CCF. But the exported bibliographic details from SOUL version 1 is not in a structured format. Tag numbers are different from MARC 21. Need more attention while making the tag numbers compatible with MARC.

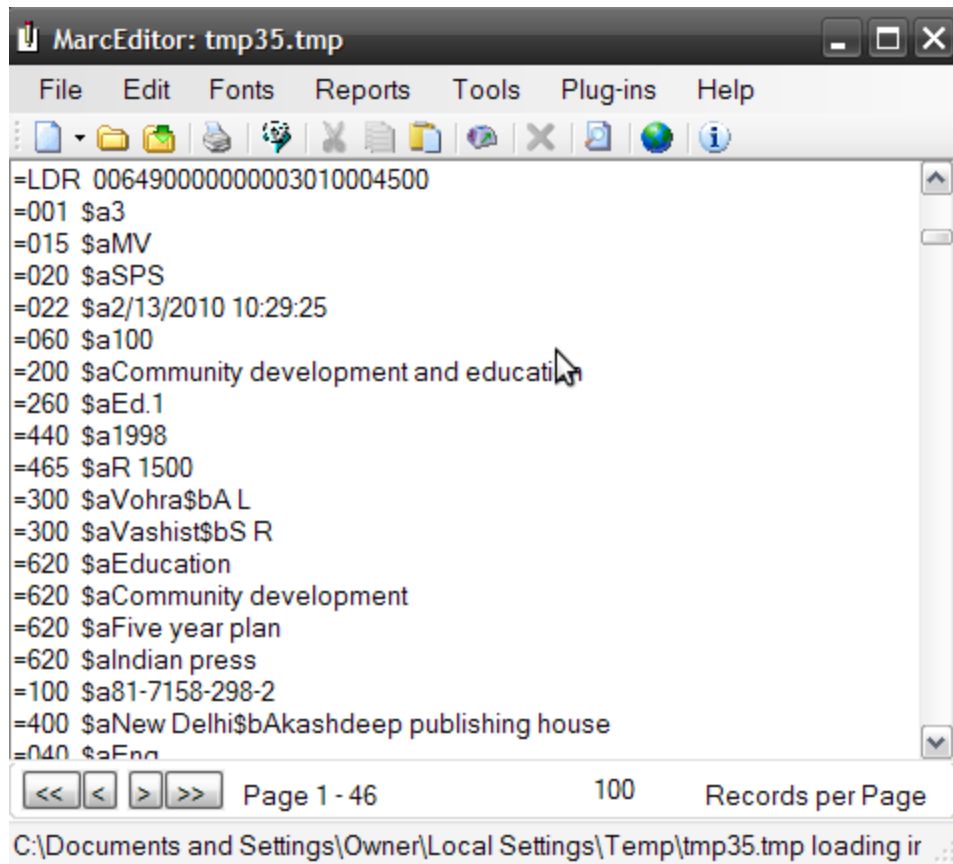


Fig. 1. View of bibliographic details from SOUL version 1.

Need to apply different tactics while trying to export data from software does not compatible with library standards. Local software developers and vendors are active in the Indian library automation market. Many libraries using library modules available with campus management software. Such systems do not care about any library-specific standards to store data. Often bibliographic details store in single or multiple tables in the database of library management systems. Bibliographic data in tabular form can be export to spreadsheet formats like CSV or Microsoft Excel. Bibliographic details store in multiple tables can be combined using the SQL query to get in single file. Graphical interface tools of databases can be used to ease the task of data export. SQL Server Express is a good choice with MS SQL. PHPMyAdmin is an excellent graphical interface tool for MySQL.

Profile of library users needs to export in spreadsheet format. Essential details like card number, name,

contact details, usernames, and passwords can be export out of the database. Transaction details like card number, accession number, issue, and due date can be port to spreadsheet file from the concerned table. The possibility of exporting transaction details can save a right amount of time in in the migration process.

Data mapping

The exported data from legacy software need to make suitable for Koha database schema. The process of preparing bibliographic, user and circulation details match with Koha requires data mapping skills. The person going to do the work should familiar with the database schema. Bibliographic details in spreadsheet or MARC format need to convert in MARC 21 or UNIMARC format. Tags for each bibliographic elements (e.g., author, title) to be matched with MARC standard. Library Management Systems using different pattern for item information; 900 for SOUL and 952 for Koha. Item information (e.g., Accession number, Price, item type) to be matched with tags devised by Koha. MARCEdit software is useful for assigning MARC tags on bibliographic details in a spreadsheet. Adding, editing, replacing options with the MARCEdit software is useful for the purpose.

Bibliographic details in non-Unicode text have to convert to Unicode format. Tools for the conversion of non-Unicode text to Unicode is available. Keep the list of bibliographic details in regional languages in a separate sheet and convert to MARC format.

Preparation of a table which indicates the differences between tags in source and destination format can reduce the errors while mapping the metadata. Tags of exported bibliographic details from SOUL version 1 has a lot of differences with MARC 21.

Table 1
Metadata tags in SOUL and Koha

Field Name	Tag Number (SOUL version 1)	Tag Number (Koha)
ISBN	100	020
Title	200	245
Edition	260	250
Author	300	100
Imprint	400	260
Physical description	460	300
Series statement	480	490
General Note	510	500
DDC Number	610	082
Keyword	620	650

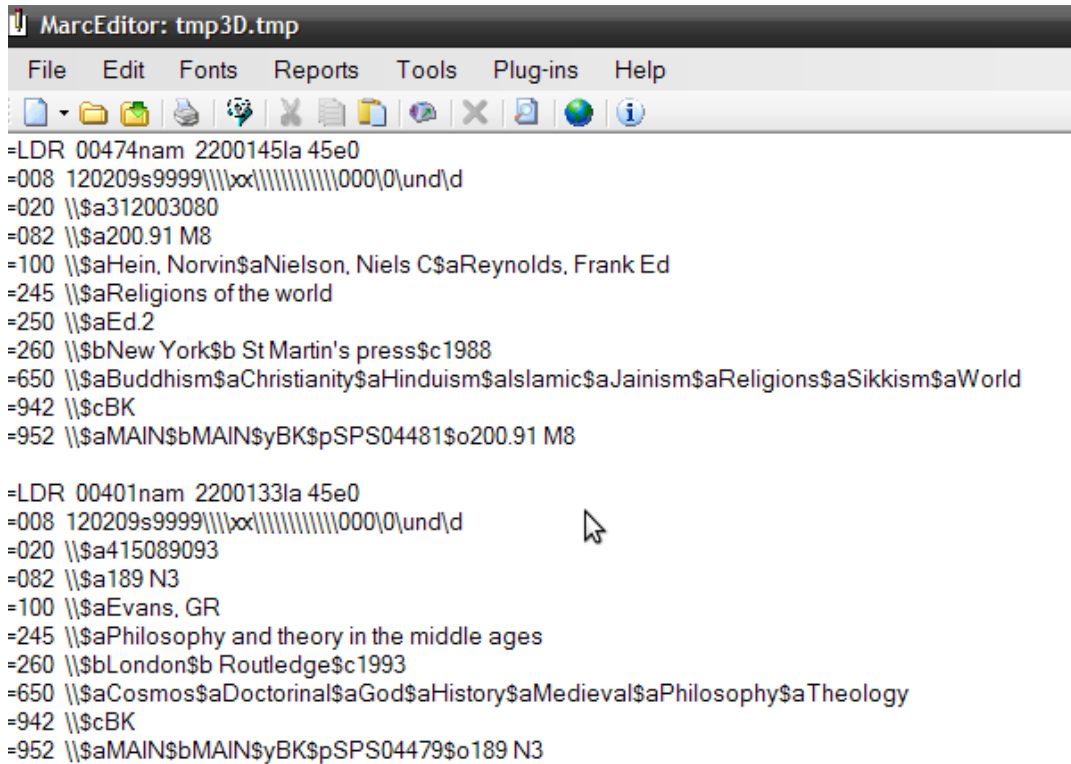
More understanding of item information part of the metadata is a crucial factor in the migration process. Certain item information missing or need modification with the legacy system can be added at the time of migration process. For example, the name of item types, shelving location, and date format can be changed with the help of MARCEdit metadata editor.

Table 2
Subfield structure of Item information in Koha.

Field name	Subfield
a	Permanent location
b	Current location
d	Date acquired
c	Shelving location
e	Source of acquisition
g	Normal purchase price

p	Accession number
y	Item type

Tag numbers and subfields numbers to be matched with destination metadata format. Unwanted characters and symbols in bibliographic details with the bibliographic details to be removed. Library professionals have to cross-check the integrity of MARC elements in bibliographic



details.

Fig. 2. Bibliographic details and item information in MARC 21 format match with Koha.

User details in a spreadsheet from legacy software prepare to match with borrowers table in Koha database. Column names of the spreadsheet should match with filed names of 'borrowers' table in Koha database. Certain Koha specific fields are mandatory in the spreadsheet. Mandatory fields like library branch code and user category require to add each user profile into the table precisely. Dates format (e.g., date of birth, date of joining) require to match with default format in Koha.

Transaction details to be devised in a spreadsheet to save it as Koha Offline Circulation file format. The format in the order of \$date issue \$cardnumber \$barcode. The provision of importing transaction details with Koha save the time of the library staff.

Data import

The newly installed Koha system need to prepare to import data from legacy automation software. Koha

checks and identifies the integrity of each record while importing data. Koha rejects the records which do not contain matching elements. For example, a library branch and item code to be the same in both Koha software and MARC file. Column names in the spreadsheet and user table should be the same. Otherwise, Koha will reject user records while importing.

Koha has the provision to import MARC records using graphical interface and command line. Importing of a small number of records using Koha staff client is a convenient option and does not require any particular skill set. The provision is very suitable for importing of a small number of records on day by day basis. Command line option is advisable for importing of a significant number of records. Bibliographic data import script with Koha is very comprehensive with many options to control over the process.

Koha has the graphical interface option to import user details in CSV format. Column names should be precise, error-free and match with field names of the table holding the user details.

CONCLUSION

Open Source library management system is a trendsetter in the library automation market. Software companies, developers and service providers encourage library software projects which encourages openness and collaboration. Such systems can handle the import and export of data from various vendor-developed systems and services including e-resources (Geller 2008)². In the present situation, libraries face difficulties with legacy systems in data migration. The success of data export depends on the nature of vendor lock, availability of library standards with legacy software and the expertise of technical support.

The attitude of library authority and staff should be in favor of accepting cutting-edge technologies for

the delivery of innovative library services. The implementation of Koha by replacing legacy systems offers a learning experience for the library staff. The library team needs to acquire new skill sets to manage the services from the innovative system. Openness and collaboration are the essences of any Open Source Software projects. Library professionals need to incorporate the sharing culture of Open Source Software to sync with the software project. Then only they can contribute back to the software community and keep in touch with the changes. A good number of prominent libraries in India have migrated to Koha. However, very few libraries have documented and shared their experiences with the community. Case studies, guidelines, and experiences on migration to Koha can act as background information and inspiration for other libraries.

References

- Balnaves, Edmund. 2008. "Open Source Library Management Systems: A Multidimensional Evaluation." *Australian Academic and Research Libraries* 39 (1): 1–13.
<https://doi.org/10.1080/00048623.2008.10721320>.
- Bhardwaj, RK, and RK Shukla. 2000. "A Practical Approach to Library Automation." *Library Progress* 20 (1): 1–9.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.629.4014&rep=rep1&type=pdf>.
- Breeding, Marshall. 2008. "Koha in Delhi, India." ALA Tech Source. Library Technology Reports, American Library Association. 2008. <https://librarytechnology.org/document/13724>.
- Chandrakar, Rajesh, and Jagdish Arora. 2009. "Library Automation in India: Past, Present, and Future." *Information Standards Quarterly* 21 (4). <https://doi.org/10.3789/isqv21n4.200901>.
- Chattopadhyay, Sougata, and Arup Sarkar. 2017. "Bibliographic Data Migration from LibSys to Koha : Experience at St. Xavier's College, Kolkata." In . Bengal Library Association, Kolkata.
<http://eprints.rclis.org/31688/>.
- Geller, M. 2008. "Managing Electronic Resources." *Library Technology Reports*.
<http://journals.ala.org/ltr/article/view/4495>.
- Karak, Sanjay, and Kuheli Dutta. 2017. "Bibliographic Data Migration in KOHA 3.18 from Existing Data Sheet: A Practical Experience." *International Research: Journal of Library & Information Science* | 7 (1): 122–32.
- Khode, Subhash, and Sunil Singh Chandel. 2015. "Adoption of Open Source Software in India." *DESIDOC*

Journal of Library and Information Technology 35 (1): 30–40.
<https://doi.org/10.14429/djlit.35.1.7897>.

Kumar, TK Gireesh, and M. Jayapradeep. 2015. "Perceptions of LIS Professionals on Open Source Integrated Library System and Adoptability of Koha over LibSys in India." *International Journal of Information Dissemination and Technology* 5 (2): 100.
<http://search.proquest.com/openview/d6d55f2afd6a05544f13f1613144f1f2/1?pq-origsite=gscholar>.

Kumar, V, and S Jasimudeen. 2012. "Adoption and User Perceptions of Koha Library Management System in India." <http://nopr.niscair.res.in/handle/123456789/15700>.

Machovec, George. 2014. "Consortia and Next Generation Integrated Library Systems." *Journal of Library Administration* 54 (5): 435–43. <https://doi.org/10.1080/01930826.2014.946789>.

Ramana, PV, VC Rao - DESIDOC Journal of Library & Information, and undefined 2003. n.d. "Use of Information Technology in Central University Libraries of India." *Gndec.Ac.In*. Accessed January 1, 2019. [https://www.gndec.ac.in/~librarian/sveri/dbit2302025\(1\).pdf](https://www.gndec.ac.in/~librarian/sveri/dbit2302025(1).pdf).

Sahu, Hemant Kumar, N. Nageswaran, S. N. Singh. 2005. "Plan and Management for Library Automation and Use of New Information Technology in Special Libraries." *3rd Convention PLANNER -2005, Assam Univ., Silchar, 10-11 Nov., 2005*.

Sexena, SC, and RK Srivastava. 1998. "Evaluation of Library Software Packages Available in India." *DESIDOC Journal of Library & Information Technology* 18 (5): 9–17.
<http://publications.drdo.gov.in/ojs/index.php/djlit/article/view/3409>.

Singh, Anil. 2003. "Library Automation and Networking Software in India: An Overview." *Information Development* 19 (1): 51–56. <https://doi.org/10.1177/026666690301900109>.

Ukachi, Ngozi B., Victor N. Nwachukwu, and Uloma D. Onuoha. 2014. "Library Automation and Use of Open Source Software to Maximize Library Effectiveness." *Library Progress (International)*.
<https://doi.org/10.5958/2320-317X.2014.00002.6>.