University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

A Functional Pathway of Bibliographic Data Migration from ADLIB to KOHA Library Management System: The MUHAS Experience

[1]Sydney Enock Msonde
Muhumbili University of Health and Allied Sciences
Email: semsonde@gmail.com

Benard E. Sengo
Muhumbili University of Health and Allied Sciences
Email: sengobenard@gmail.com

Charles V. Mwalyego
Muhumbili University of Health and Allied Sciences
Email: vcmwalyego@gmail.com

William J. Mviombo
Muhumbili University of Health and Allied Sciences
Email: mviombojulius@gmail.com

Abstract

The early dominance of proprietary library management software (LMS) has led higher education institutions in developing countries to automate their library operations and services. However, high prices, stringent license agreements, limited access to training, and a lack of control over data necessitated that many higher education institutional libraries migrate their bibliographic data from proprietary software to open-source library management systems. This research aimed to demonstrate the functional pathways involved in bibliographic data migration from ADLIB, a proprietary library management system, to KOHA, an open-source LMS. This research was an experimental design study focusing on data migration and systems integration. A committed group of academic librarians with extensive experience in both data migration and systems integration completed the data migration task. The data for this study were extracted from the ADLIB system existing at Muhimbili University of Health and Allied Sciences (MUHAS) and during multiple tests of importing bibliographic records into the KOHA system. For eight months, the MUHAS academic librarians migrated approximately 63,000 bibliographic records using a step-by-step data migration pathway. Findings from this study provide insight for other higher education institutional libraries in developing countries that aspire to transition successfully from the ADLIB to the KOHA system. The paper highlights the successes and shortcomings of the bibliographic data migration process experienced at MUHAS, and how librarians would address the identified constraints for effective data migration and the provision of modern library services.

Keywords: Library management system, Data Migration, Bibliographic data migration, KOHA system, ADLIB system

https://dx.doi.org/10.4314/udslj.v20i1.6

[1] Corresponding Author

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

Introduction

Library Management Systems (LMS) are software solutions designed to automate core library functions such as cataloguing, circulation, acquisitions, and service delivery (Breeding, 2018). Unlike open-sources, proprietary LMS have been developed and dominate the market due to their reliability, advanced features, and vendor support (Reddy & Kumar, 2013). Globally, academic libraries have widely adopted systems such as Ex Libris Alma for its cloud-based infrastructure and analytics, SirsiDynix Symphony for its scalability in public and university libraries, and Innovative Interfaces' Sierra for its robust cataloguing and discovery tools (Breeding, 2018). In India, higher learning institutions use open-sources (e.g, KOHA) alongside proprietary LMS for hybrid functionality (Pankaja & Mukund Raj, 2013), while in South Africa, universities rely on proprietary systems (e.g, Aleph) to manage extensive digital collections (Masenya & Ngulube, 2021). Similarly, the University of Nairobi in Kenya employs Libsys to streamline workflows (Otike & Barát, 2021). These systems typically offer modules for Online Public Access Catalogue (OPAC), RFID integration, interlibrary loan management, and compliance with metadata standards (Chowdhury & Chowdhury, 2019). In Tanzania, the early influence of proprietary LMS like ADLIB spurred automation in institutions such as the University of Dar es Salaam (Samzugi, 2016). A good example of proprietary LMS used in most higher learning institutions in Tanzania is ADLIB software, which has been in the market for over 40 years. This software played a crucial role in managing the library's information resources, enabling the effective and efficient provision of library services to patrons. In Tanzania, no universally dominant LMS is used among higher learning institutional libraries, as each has its own. The commonly used LMS in Tanzania include ADLIB, ABCD, and E-print, proprietary software. A significant number of libraries received financial support from SIDA-SAREC and were able to subscribe to ADLIB licenses for decades (Samzugi, 2016).

The ADLIB offers various benefits, including integrated management of diverse collections, customizable metadata fields, and support for international cataloguing standards. It also ensures consistent cataloguing through authority control and supports digital object management, enhancing access to digital collections. Despite the positive benefits that ADLIB offers, many higher education institutional libraries have begun promoting open-source software, such as KOHA, for managing library operations. Several factors, including stringent licensing agreements and limited access to training, may have contributed to this situation (Todd, 2018), as well as a lack of control over data when ADLIB software is used (Kumar & Majeed, 2019). Similarly, inadequate funds to purchase or maintain proprietary software, as well as existing university ICT policies that encourage the use of open-source software, contribute to this movement. These factors necessitated that many higher-learning institutional libraries migrate their bibliographic data from proprietary software to open-source LMSs, such as.KOHA (Kumar & Majeed, 2019). Earlier generations of LMS, such as ADLIB software, were developed to meet the needs of the printed collection. In contemporary library services provision, it is essential to automate libraries with integrated library management systems (ILMS) that cater to both print and digital resources. Various open-source ILMS exist in the fields of library and information science. These include BiblioteQ, NewGenLib, OpenBiblio, CodeAchi, and KOHA, among others. Although all these provide proper management of the library information resources, KOHA software demonstrated

comprehensive features and functionalities that meet the needs of any library (Machovec, 2014). In that regard, institutional libraries with no ILMS may wish to automate library operations and services directly using KOHA from scratch (Ahmed, 2024). On the other hand, Tella and Oladeji (2017) argue that libraries with limited funds to cater to high-priced proprietary licenses, such as ADLIB, may opt to move their bibliographic records to KOHA, an integrated library management system.

Muhimbili University of Health and Allied Sciences (MUHAS) automated its library operations and services using ADLIB in 2005, when it was a constituent college of the University of Dar es Salaam and was known as MUCHS at that time. The annual subscription fee for the ADLIB license was covered through the SIDA/SAREC project until 2015, when such financial support was terminated. Since then, there has been no ADLIB system upgrade, limited training on new software updates, and the security of data has remained uncertain. Such a situation, following the university's ICT policy, which promotes the use of open-source software (Samzugi, 2016), necessitated the MUHAS library's migration of its bibliographic records from ADLIB to KOHA software. The literature demonstrates that KOHA software has provisions and the capability to work with add-on modules (Kumar & Majeed, 2019), which can extend library operations and services. Therefore, MUHAS found that the KOHA software had great promise for the effective and efficient provision of modern library services.

Previous case studies of this nature reported that bibliographic data migration relied heavily on institutional IT departments (Kohn and McCloy, 2010) and on third-party contractors (Genoese and Keith, 2011). In the current case study, the bibliographic data migration from ADLIB to KOHA software was accomplished by a dedicated team of academic librarians with experience in systems integration and customisation. Their experience was demonstrated through the successful mapping of metadata fields and customisation of KOHA to accommodate existing data structures. They also ensured seamless data transfer with minimal loss or errors, preserving the integrity and accessibility of records in the new system. As such, these library professionals were familiar with the required data structure, mapping, and various phases of the migration process. This phenomenon was contrary to Todd's (2018) The study reports that the data migration team at Northern Marianas College consisted of librarians with no prior experience in data migration or systems integration. In the current case study, the bibliographic data migration process encountered several constraints, including the presence of data in an unstructured format and vendor lock-in. Despite these shortfalls, the MUHAS data migration team completed the bibliographic data migration from ADLIB to KOHA without any external support. As such, the step-wise data migration and system troubleshooting improved academic librarians' database skills and increased familiarity with the new system.

Problem statement

The MUHAS Library was using the ADLIB system, which did not comply with various library information exchange standards, such as MARC, Z39.50, Dublin Core, and UNIMARC, making it challenging to communicate with other library systems worldwide. Such system inefficiency impairs normal library operations, including cataloguing and classification, interlibrary loan, and the circulation of library resources. Furthermore, ADLIB, being a commercial library system, required regular system upgrades to accommodate new features and capabilities while also improving system security. Such practice has significant financial ramifications, which most higher learning institutions in developing countries, such

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

as MUHAS in Tanzania, could not afford. As a result, it was required to migrate bibliographic data from a commercial (ADLIB) to an open-source system (KOHA). Therefore, the purpose of this study was to demonstrate a functional pathway of bibliographic data migration from ADLIB to the KOHA library management system.

Study objectives

The overall objective of this case study was to demonstrate the functional pathways of bibliographic data migration from ADLIB to KOHA software. The specific objectives were threefold:

1. To demonstrate the step-by-step pathways for successful bibliographic data migration from ADLIB to KOHA open-source software.
2. To evaluate impending shortfalls during the bibliographic data migration process.
3. To recommend strategies to ameliorate the earmarked constraints for an effective data migration process.

Review of related literature

The literature was reviewed in accordance with the stated research objectives. This section highlights the concept of bibliographic data migration, including experiences with data migration from various systems, as well as the factors that influence the data migration process in academic libraries. Subsequent sections elaborate on each of these items.

The concept of bibliographic data migration

Moving data from one system to another is known as data migration. This process may be broken down into two parts: (a) extracting data from the source system into an extracted file, and (b) loading that extracted file into the destination system (Todd, 2018). However, transforming data from one library management system to another appears to be a complex process that requires specialised skills and expertise to navigate numerous difficulties. For example, some literature indicates that users of LIS, such as Libsys and ADLIB, were not granted access to the source code, and front-end programs were the only tools available for manipulating the data (Das, 2014; Vimal Kumar & Majeed, 2019). Due to the unstructured output from the source file and the absence of important data fields, such as title, author, and accession number, some records may be lost entirely or partially during the data conversion process.

Some of the best practices and strategies for data migration include assessment of the quality and integrity of the source data. This step involves performing a data audit and analysis to identify any data errors, anomalies, duplicates, or gaps that must be resolved before the migration (Anyaogu, 2019). improve the accuracy, completeness, and consistency of the data (Hussein, 2021). It is also important to design and test the data migration plan. This step involves choosing the appropriate data migration method, such as Big Bang or Phased, and the data migration tools, such as ETL (Extract, Transform, Load) or ELT (Extract, Load, Transform) (Das, 2014; Todd, 2018; Mushi, Mwantimwa and Wema, 2020). This involves mapping the data fields and transformations between the source and target systems, as well as defining the data migration rules and policies. The data migration plan should be tested in a simulated environment to verify its functionality, performance, and reliability.

Another issue concerns the execution and monitoring of the data migration process**.** This involves following the data migration plan and schedule and ensuring that the data is securely and efficiently transferred from the source to the target system. The data migration process should be monitored and controlled using data quality metrics, logs, and reports, and any issues or errors should be promptly resolved or escalated (Thalheim & Wang, 2013). On the other hand, validation and evaluation of the data migration results should be conducted to ensure the accuracy, completeness, and consistency of the transferred data. This involves verifying that the data in the target system is accurate, complete, and consistent with the source data, and that it meets the business requirements and expectations. Data quality checks, audits, and feedback can help validate and evaluate the data migration results and identify any areas for improvement or optimisation (Thalheim & Wang, 2013; Hussein, 2021). In general, migrating bibliographic data is a crucial procedure that calls for meticulous planning, close attention to detail, and cooperation between library employees (Das, 2014). It ensures the seamless transfer of critical library data to new systems, benefiting both users and librarians.

Data migration experiences from various systems

Data migration typically involves mapping data to new systems and decommissioning legacy systems. Moving on-premises data systems and infrastructure to cloud-based settings is becoming a typical use case, as more businesses invest large sums of money to reduce storage costs and increase productivity (Todd, 2018; Kumar & Majeed, 2019). Because of the process's complexity and limitations, only around 16per cent of data migration projects are completed within the allocated time and budget, according to survey results from various industries (Potter, 2007).

Considering the growing shift towards digital infrastructure, libraries have increasingly embraced Information and Communications Technology (ICT) to enhance the management of their information resources and streamline operations (Todd, 2018; Innocent & Masue, 2020). Central to this transformation is the adoption of Integrated Library Systems (ILS), both commercial and open-source, which serve as vital tools in modern librarianship. These systems facilitate the automation of library functions and align with broader institutional trends of migrating to cloud-based platforms for improved efficiency, cost reduction, and service delivery. As with other data-intensive sectors, implementing and transitioning between ILS platforms often involves complex data migration processes that must be carefully managed to ensure continuity and performance (Anyaogu, 2019; Samzugi, 2016)

Open-source library management systems have recently made a footprint in the automation software market (Vimal Kumar & Majeed, 2019). KOHA software, being one of the open-source LMSs, is becoming popular among academic libraries due to the availability of comprehensive functional modules and an active community (Kumar, 2015; Todd, 2018). Some libraries initially automated their resources and services using commercial software, such as ADLIB, and are now opting to migrate their bibliographic data from their old systems into KOHA. The primary reason for this option is the financial inability of academic libraries to afford the high cost of proprietary licenses. Similarly, university ICT policies promoting open-source software were another factor that necessitated academic libraries to migrate the bibliographic records from ADLIB to KOHA software (Samzugi, 2016). The current study aligns with these broader trends by examining the bibliographic data migration process from ADLIB to KOHA at MUHAS. It highlights the practical challenges, institutional motivations,

and technical strategies involved in transitioning from a proprietary system to an open-source platform, thereby contributing to the growing knowledge on data migration experiences across different library systems.

Factors influencing the data migration process

For libraries requiring technical assistance, the transition to Koha is relatively straightforward. KOHA provides a range of conveniences for individuals seeking data migration. Data conversion to KOHA is made simple by the availability of several data import methods (including GUI and command line tools), compatibility with bibliographic standards, open database schema, documentation, and community assistance. The skill sets required for an individual handling data migration include proficiency in workflow, database management, and operating systems based on Linux. A study by Kumar and Majeed(2019) regarding data migration from the Legacy system to KOHA, the following factors may influence data migration:

- Data Volume: Considering the volume of data is essential as it directly impacts the complexity, duration, and cost of the migration process. Larger data volumes may necessitate more sophisticated tools and could increase the risk of data loss or corruption, making it vital to plan for data backup and recovery.

- Type of Workloads: It is essential to consider the type of workload that the new system will handle, such as Virtual Machines (VMs), backups, and databases. This guides the performance, storage, and capacity planning of the new system, affirming its readiness to handle these workloads efficiently post-migration.

- Speed to Completion: The expected timeline or speed to completion of the migration process is a significant factor. A shorter timeline may require more resources or advanced tools, emphasising the need to balance speed with accuracy and efficiency in the migration plan.

- Regulatory Compliance: Prioritising compliance with all relevant data protection and privacy laws is essential when migrating sensitive or regulated information. Non-compliance can result in legal penalties and loss of customer trust, making it a crucial factor to consider in the migration process.

- Project Resources: Assessing the available resources is vital for the success of the migration project. The team's skills and expertise, the capabilities of the migration tools, and the project budget can all influence the migration strategy and its success.

On the other hand, Kumar and Majeed (2019) highlighted additional factors to consider during the migration process, including data quality, business continuity, user training, and data complexity, to ensure a smooth and successful exercise. It is also crucial to remember that successful data migration requires meticulous planning, a thorough understanding of system nuances, and adherence to best practices. Therefore, by considering these factors, organisations, including academic libraries in Tanzania, can achieve a smooth and successful transition of their valued data from one system to another system of their current interest (Potter, 2007; Karak & Dutta, 2017; Todd, 2018). In the current study, these factors played a significant role in guiding the migration process at MUHAS. The decision to shift from ADLIB to KOHA was influenced by financial constraints and institutional policies that favoured open-source solutions. These considerations informed the planning and execution

phases, during which a careful evaluation of data quality, system compatibility, and staff readiness was conducted. Moreover, the use of a skilled team with experience in system integration and the application of standard data validation techniques ensured that the migration aligned with best practices drawn from previous studies and experiences.

Adoption of LMS in Tanzanian Higher Learning Institutions

Global interest in library automation has grown, particularly in emerging nations such as Tanzania, where the use of ICT has expanded dramatically across various domains of life. In the age of the knowledge society, ICT has a profound impact on almost every endeavour undertaken by humans. In a similar vein, libraries are using ICTs to create databases and automate technical and administrative procedures. This incredible rise in ICT has increasingly helped Tanzania's university libraries to modify the services they provide. In Tanzania's university libraries, there has been a transition from manual to automated systems to enhance efficiency, quality, and operations when providing learners with support services to a broader user base.

Tanzanian universities have been utilising a range of library administration software programs for information access, distribution, archiving, and utilisation, as well as for storing bibliographic records and administrative tasks (Karak & Dutta, 2017). These libraries used a variety of automation tools (software), including ADLIB, KOHA, DSpace, E-Print and ABCD (Samzugi, 2016). On the other hand, Samzugi's (2016) study findings reveal that there was no universally available dominant library management software among university libraries in Tanzania, as every university had its own (Samzugi, 2016). Given this varied landscape of library software usage in Tanzanian universities, the role of Library Management Information Systems (LMIS) becomes particularly significant. These computer-based applications are designed to support efficient management of library functions such as cataloguing, circulation, acquisitions, and user management (Todd, 2018; Vimal Kumar & Abdul Majeed, 2019). They enable libraries to streamline operations, manage bibliographic data, and provide access to information resources across different physical and digital platforms (Sahoo & Saikia, 2019). However, the absence of a standardised system across institutions, as noted by Samzugi (2016), suggests not only a fragmented implementation approach but also underscores the need for coherent strategies in LMIS adoption.

The adoption of ILMS has not been an easy path in most universities in Tanzania. The transition to using such technology has been associated with several challenges, including poor infrastructure, financial constraints, inadequate support, a lack of or insufficient knowledge, rapidly changing technology, and resistance to change among professionals (Innocent &Masue, 2020; Mushi & Mwantimwa). These challenges can be categorised into institutional and human. The most common institutional factor is poor infrastructure related to communication. Sources of power supply, computer laboratories, as well as the ICT technical support unit and a lack of poor library policies that set milestones in place. On the other hand, human factors have been associated with barriers to LMIS adoption in HLIs, in a way that library professionals' understanding can also impact the effective use of such technology in the library setting. This inconsistency reflects broader organisational and human challenges that shape how libraries leverage technology to fulfil their information management objectives.

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

The extant literature shows that if information professionals (Librarians and IT personnel) do not understand the meaning and impact brought about by the ILS, then they are likely to resist or avoid using it, resulting in institutional failure to adopt it (Samzugi, 2016). Consequently, these challenges have led to the failure of academic libraries in Tanzania to fully adopt and integrate librarianship as an innovative approach to enhance library service delivery (Mushi, Mwantimwa, et al.). Similarly, (2016) affirms that, in addressing financial challenges, most library automation projects in academic libraries in Tanzania rely on donor support as the major source of funding for library automation in public university libraries. On the other hand, Samzugi (2016) found, in a study assessing the status of library automation in Tanzanian libraries, that there is no universally available, dominant library management software among public university libraries in the country. As such, there was an urgent need to develop standard software among public university libraries to facilitate information sharing and exchange among these institutions. In addition, barriers to effective library automation ought to be overcome to ensure the sustainability of library automation (Das, 2014; Samzugi, 2016; Innocent and Masue, 2020).

Mulimila (2000) conducted a regional study on Information Technology (IT) applications in East African government-owned university libraries for the 1987-1997 period, which only covered two state-owned university libraries in Tanzania. These were the University of Dar es Salaam and Sokoine University of Agriculture libraries. Among the East African university libraries surveyed, only Moi University Library had automated its catalogue and circulation at that time. The situation of the two public-owned university libraries in Tanzania has since changed as they have also jumped onto the library automation bandwagon. Moreover, Mulimila's (2000) study was carried out many years ago. Planning for the automation of the University of Dar es Salaam library has been the main focus of the study by Manda, (2003) & Wamunza (2003). Although this corpus of work has yielded valuable insights into the process of library automation, the two studies do not adequately demonstrate the outcomes that follow automation. This statement means that there is a gap in understanding the real-world effects or benefits that libraries experience once automation is completed.

On the other hand, a study on the information-seeking behaviour of library users in a changing library environment using the Faculty of Law staff members of the University of Dar es Salaam as a case study has been conducted (Felicia, 2015). The study identified significant changes, including the computerisation of library services. The scope of this study was limited to the University of Dar es Salaam and, therefore, could not provide insight into the library automation of other public universities. Kasulwa's (2008) report mentions the automation efforts in Tanzania, citing the University of Dar es Salaam library as an example, and concludes that some universities were in the process of automating their respective libraries. Little else, however, is known about the library automation status of other universities.

A study by (2016), focusing on the status of Library automation in Tanzania's public universities, found that public university libraries utilise various types of software, including both open-source and licensed software, to automate their library operations and services. The five types of software used include ADLIB, KOHA, DSPACE, E-Print, and ABCD. The study findings reveal that there was no universally available dominant library management software among public university libraries in Tanzania, as every university had its own. These findings align with several other studies, including those by Malik (1996), Muneja (2010), and

COSTECH (2012). Overall, these studies affirm that standard library software was non-existent. This was a significant issue of concern that hindered resource sharing.

Research design and methods

This study employed an experimental design involving a dedicated team of academic librarians with extensive experience in data migration and systems integration. The migration process from ADLIB to KOHA was our independent variable, while data integrity, system performance, and user satisfaction were the key dependent variables. The study employed a phased (incremental) data migration approach, consisting of eight phases. This research used Marc Edit software for extraction and data processing. Similarly, a spreadsheet program uses CSV and Excel file formats to clean and analyse data. A detailed explanation of all the data migration pathways from ADLIB to KOHA, an open-source software, is presented in the subsequent sections.

Data migration pathways

The decision to migrate the bibliographic data from the ADLIB System to the KOHA open-source system at MUHAS Library was a challenging one. However, this became possible given the number of librarians with technical expertise in ADLIB and KOHA Systems at MUHAS. A technical team was formed to study both KOHA and ADLIB systems. KOHA offers various services and functionalities for data migration. KOHA offers two primary data migration options: the Graphical User Interface and the command-line option. Several issues were considered before and during bibliographic data migration from ADLIB to KOHA. These issues include compatibility with bibliographic standards, open database schema, documentation, and community support. These issues have made the data migration process at MUHAS Library more straightforward. During the data migration process at MUHAS Library, the data was migrated in eight phases from ADLIB to the KOHA system (see Figure 1). These phases include assessing the systems and planning, extracting data from ADLIB, and transforming and cleaning the data. Other phases included data matching, mapping, and importing into KOHA, as well as system Testing. Each of the bibliographic data migration phases has been described in subsequent sections.

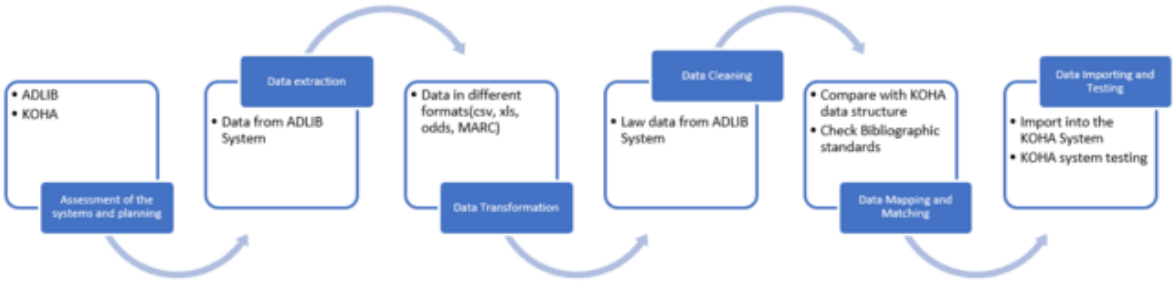

Figure 1: Steps involved in bibliographic data migration

Source: Researchers Creation, 2024

Assessment of the Systems and Planning

The aim of assessing the system was to understand the data structure and format of the ADLIB system and the KOHA integrated library management system, in order to examine their compatibility. The data fields for both ADLIB and KOHA systems were examined for compatibility, enabling the team to conduct potential data mapping smoothly. Table 1 shows

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

data fields used in the KOHA system. The corresponding data fields for the ADLIB system are shown in Figure 2:

Table 1: Data Fields in KOHA Systems

Title Statement	**Doing qualitative research: a practical handbook**
Personal Name	Silverman, David
Edition Statement	3rd ed.
Place of publication, Distribution	London
Name of Publisher, Distributor	Sage
Date of publication, Distribution	2010
Extent	xiv,456p.
Other Physical Details	
Dimensions	26cm.
Type of Unit	Book
ISBN	9781848660034
Barcode	000000102814

A deeper examination of these systems reveals that both ADLIB and KOHA systems had similar data fields, as shown in Table 1 and Figure 2. Therefore, this similarity made the data extraction process possible. The only difference observed between these systems was the presence of the field of copy number and barcode number. Such that a copy number in ADLIB was regarded as a barcode number entry in KOHA, this difference created the likelihood of errors during the data migration process.

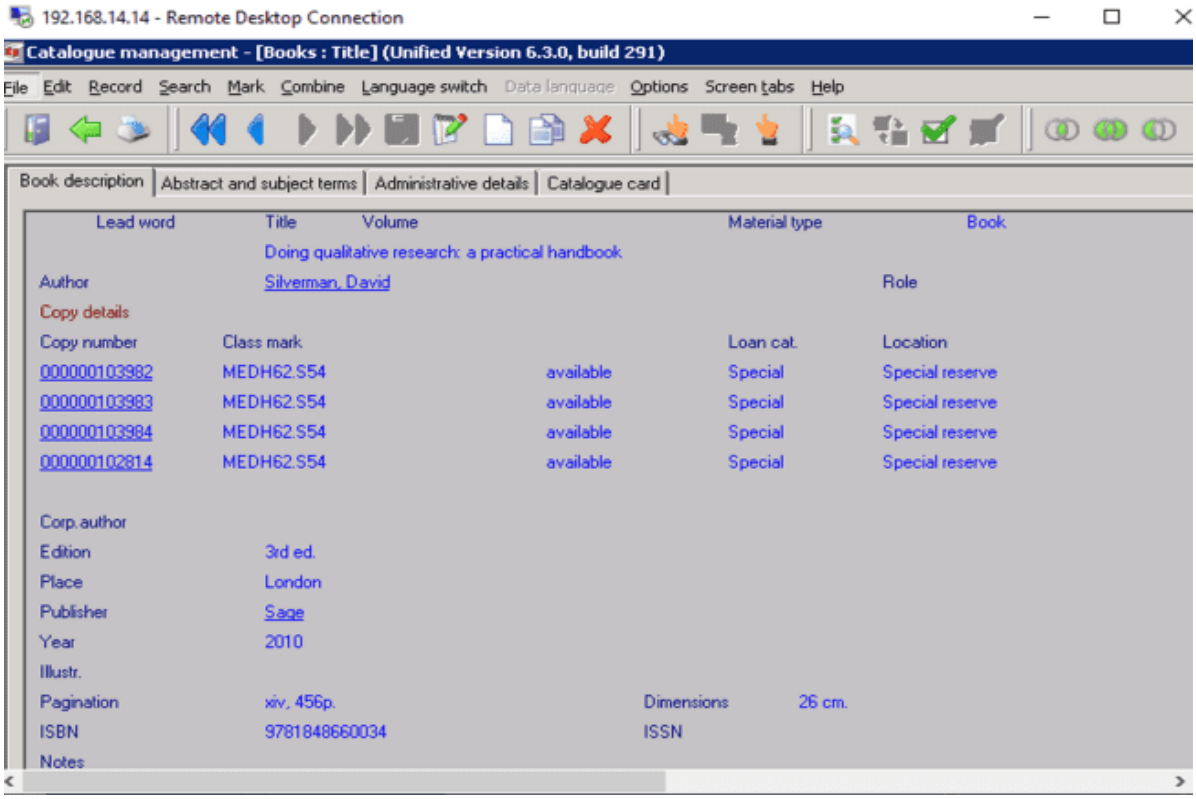

Figure 2: Data Fields in ADLIB

Data Extraction from ADLIB

Bibliographic data extraction can be defined as the process of locating, identifying, and retrieving relevant data, followed by preparing it for the transformation stage. This step is crucial for integrating various types of data and preparing them for further processing. In the context of this study, the extraction of bibliographic data from the ADLIB system was successfully carried out in a standardised and structured format. The process began with the identification of the relevant bibliographic records to be migrated. Using ADLIB's built-in export functionality, data were exported in a format compatible with recognised standards such as MARC (Machine-Readable Cataloguing) and CCF (Common Communication Format). The system's compliance with these standards facilitated smooth extraction and ensured interoperability with the target system (KOHA).

Data were then exported using the MARC editor tool, which enabled the team to review and validate the metadata fields before proceeding. During this step, fields such as title, author, publication date, subject headings, and classification numbers were extracted with their corresponding MARC tags. The exported files were saved in a format suitable for transformation and later import into KOHA, such as MARCXML or .mrc. This structured approach ensured that all critical bibliographic information was accurately extracted, preserved, and formalised in the required file format, thereby minimising data loss and preserving metadata integrity. Figure 3 presents a sample of the data structure extracted from ADLIB using the MARC editor.

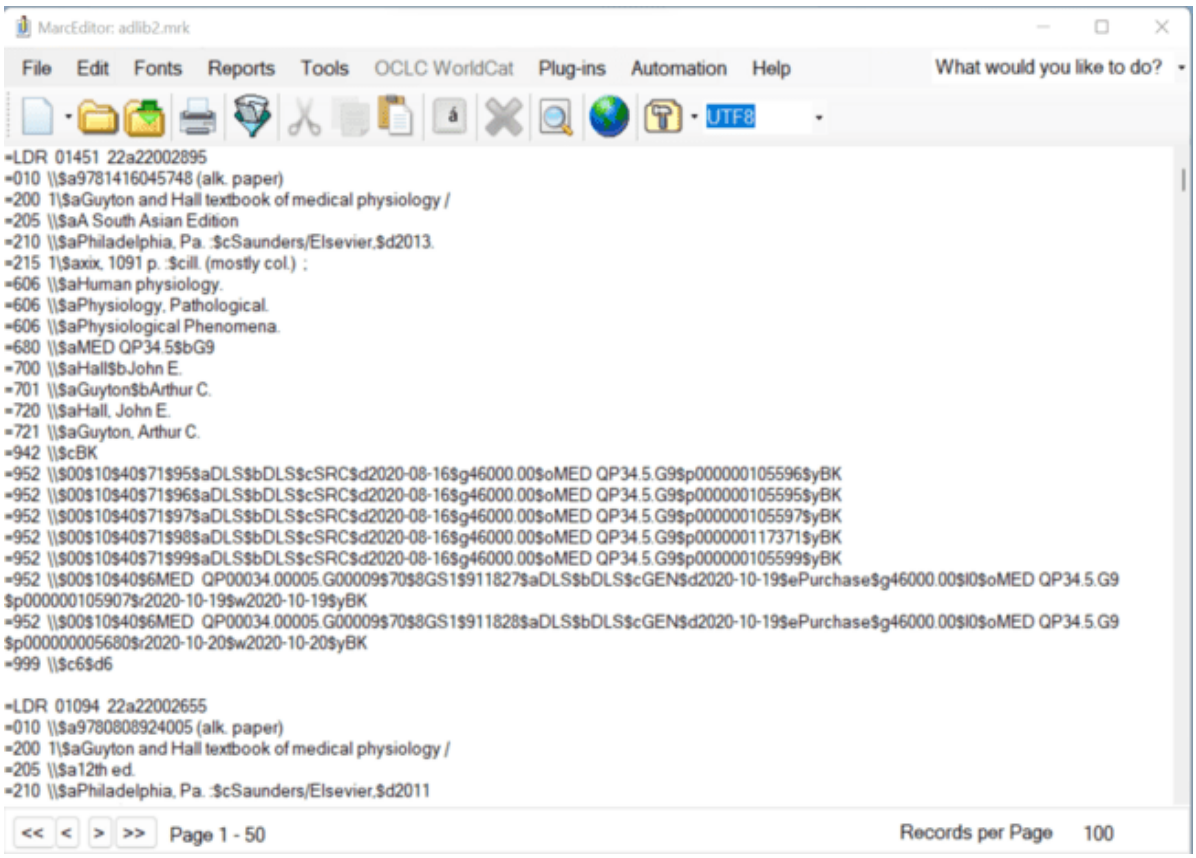

Figure 3: Data Structure Extracted from ADLIB

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

Some of the information extracted from the ADLIB system included titles, authors, subjects, publishers, publication dates, and other metadata. The extracted file was converted into a CSV file format to facilitate data transformation and cleaning. A total of 32,313 bibliographic data points were extracted from the ADLIB system, ready for data transformation and analysis. The level of data extraction was equivalent to 100 per cent.

Data Transformation

Once the bibliographic data had been successfully extracted from the ADLIB system, it was ready for refinement. During the transformation phase, data were sorted, organised, and cleaned. For example, the copy number field in ADLIB was replaced with a barcode number to match the data field in KOHA. Similarly, duplicate entries were deleted, missing values removed or enriched, and audits were performed to produce data that is reliable, consistent, and usable. In other words, this phase aimed to convert the extracted data into a format compatible with KOHA (i.e. the map fields from ADLIB to match with corresponding fields in KOHA). Figure 4 provides insight into the data structure transformed from ADLIB to the KOHA system.

While the primary focus of the data transformation was on structural alignment—such as replacing the *copy number* field with *barcode number*, removing duplicates, and enriching missing values—it was also important to assess the underlying bibliographic frameworks employed by both ADLIB and KOHA. The ADLIB typically uses formats like MARC 21 or custom metadata structures, depending on the institution's configuration. KOHA, on the other hand, natively supports and relies heavily on MARC 21, particularly for managing bibliographic and authority data.

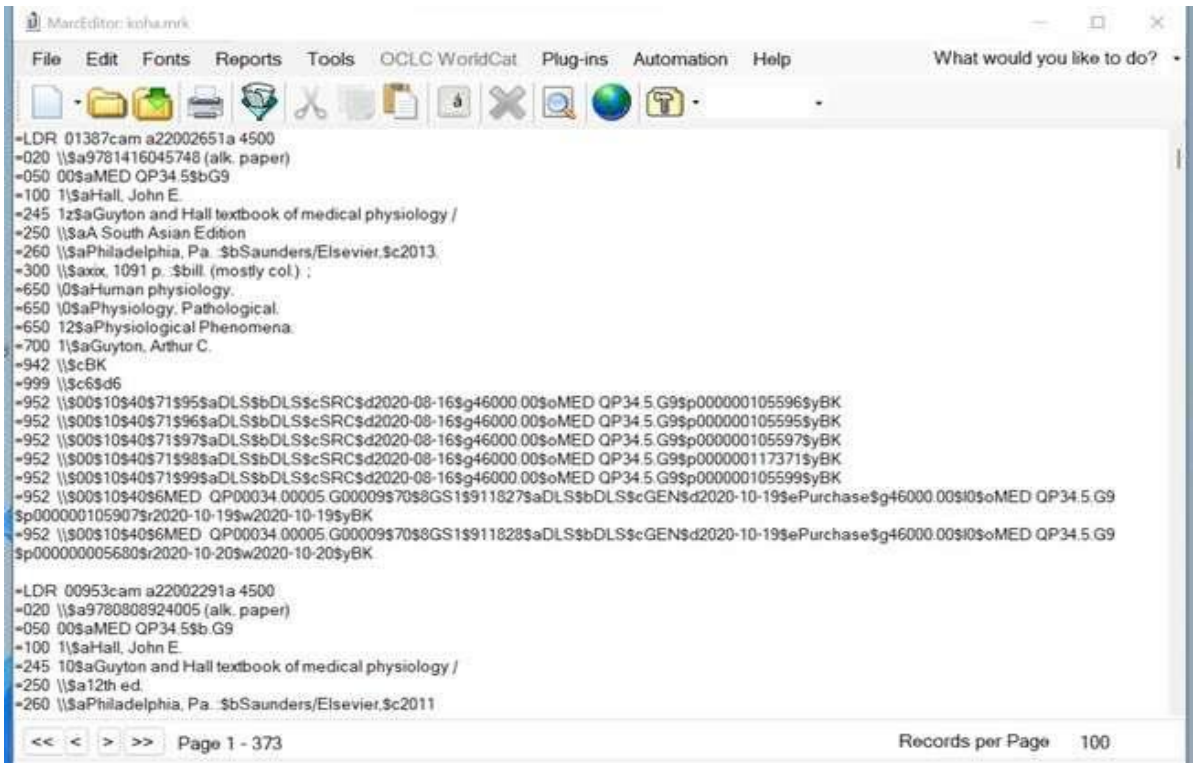

Figure 4: Data Structure Transformed from ADLIB to KOHA

As part of the data mapping and transformation process, efforts were made to ensure that the bibliographic records adhered to MARC 21 standards where applicable. This included aligning key fields such as title, author, subject headings, and item-specific data to ensure compliance and usability within KOHA's MARC-based cataloguing environment. Therefore, while the focus was mainly on field mapping and structural compatibility, attention was also paid to the compatibility of bibliographic standards, ensuring the integrity and interoperability of migrated data within KOHA's MARC-compliant framework, which facilitated data cleaning.

Data cleaning and Validation

By definition, data cleaning involves the process of correcting or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete bibliographic records within a given system (Ahuja, Singh, & Simon, 2024). In this study, the researchers conducted bibliographic data cleaning to ensure the extracted data was clean and removed inconsistencies, errors, and duplicate records. The MarcEdit software (i.e., MARC-agnostic) was used to create, edit, and manipulate library metadata in various MARC formats and standards. Therefore, some errors and duplicates were removed to ensure the accuracy and completeness of the data. Figures 5a–5c illustrate the stepwise data cleaning and validation process.

Remove duplication

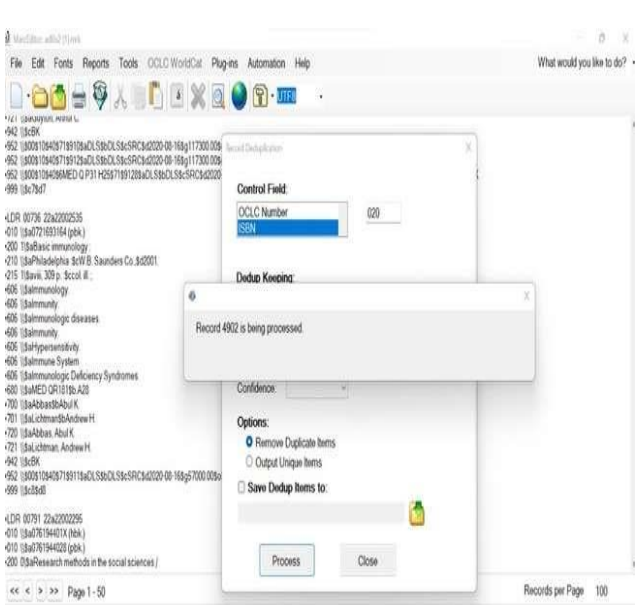

Fig. 5a

Remove empty fields or subfields

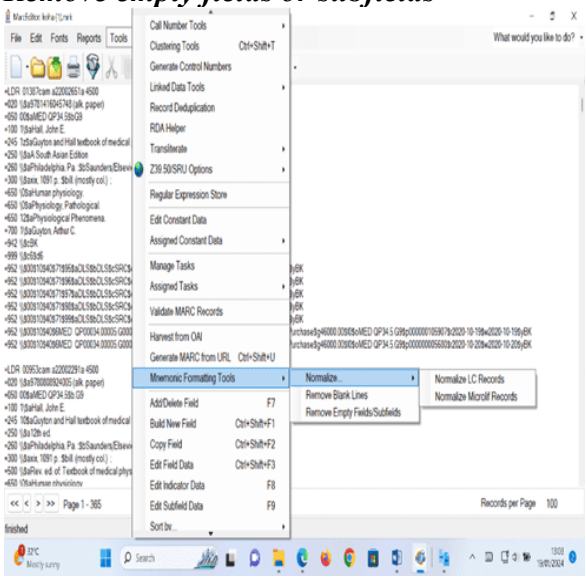

Fig. 5b

Validation of Records

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

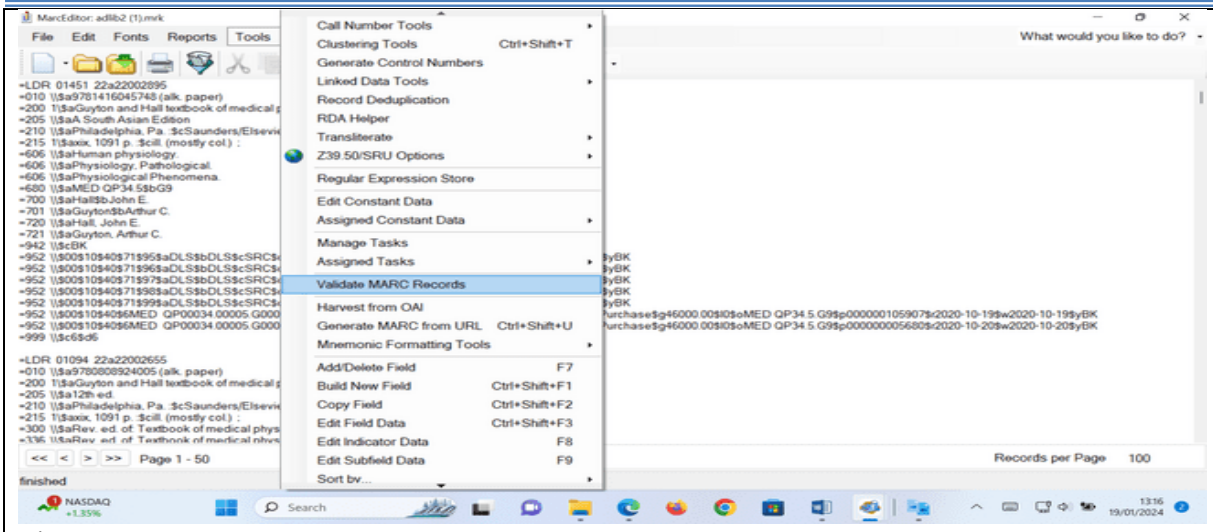

Fig. 5c

This study used the Marc Edit tool to identify duplicate records based on ISBN and the unique identifier. Determination criteria for merging or eliminating duplicates were applied. The criteria involved include preferring records with more complete information or selecting the most recently updated record. The empty fields and subfields were removed to ensure data consistency and cleanliness, resulting in normalised data. Thereafter, the data were checked and validated to meet KOHA`s standards.

Data mapping and matching

Immediately after completing data cleaning and validation, the output file was prepared for importing into the KOHA System. The matching and mapping of tag numbers between the ADLIB and KOHA systems was made to have a clear and clean file during importation. Table 2 shows the Tags of bibliographic details of ADLIB that were matched with the tags of KOHA bibliographic details.

Table 2: Matched tags of ADLIB and KOHA bibliographic records

Field name	Tag number for ADLIB	Tag number for KOHA
ISBN	100	020
Title and statement of responsibility	200	245
Edition	260	250
Author	300	100
Physical Distribution	460	300
Series Statement	480	490
General note	510	500
LC Number	610	082

The data mapping process ensured that fields such as title, author, ISBN, and call number in ADLIB were correctly mapped to their corresponding fields in KOHA. Consequently, data matching helped identify and merge duplicate records, ensuring that each bibliographic entry is unique and accurate. Thus, these processes enabled a seamless migration, maintained data integrity, and enhanced the overall functionality of the new system.

Data import into KOHA

The output file, after data cleaning, was converted into an .mrc file format for easy import into the KOHA System. In the staff client interface, the following steps were undertaken to import bibliographic data into KOHA: First, the Tools module was opened, and then the data were staged for import, as shown in Figures 6a to 6d. Thereafter, the .mrc file was uploaded to initiate the importation process. See figures 6c and 6d.

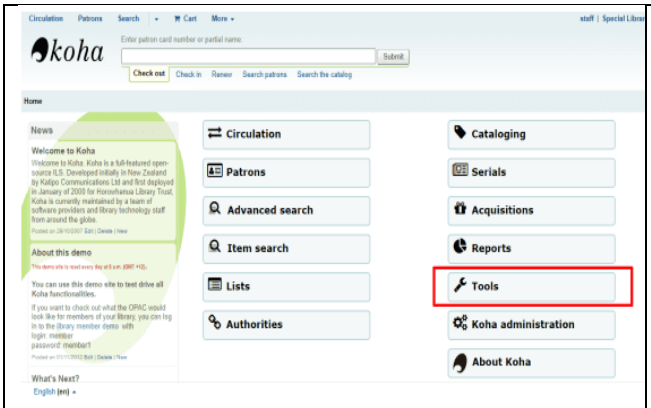

Figure 6a:

Figure 6b:

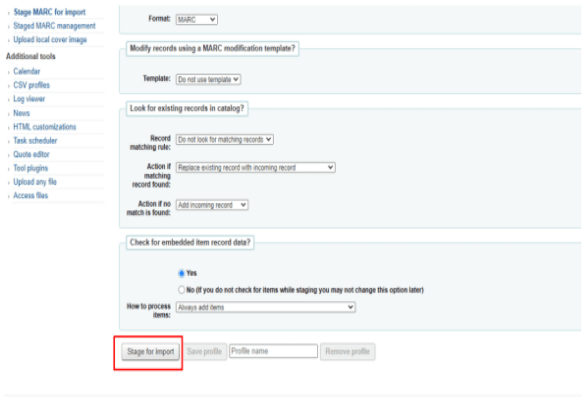

Figure 6c:

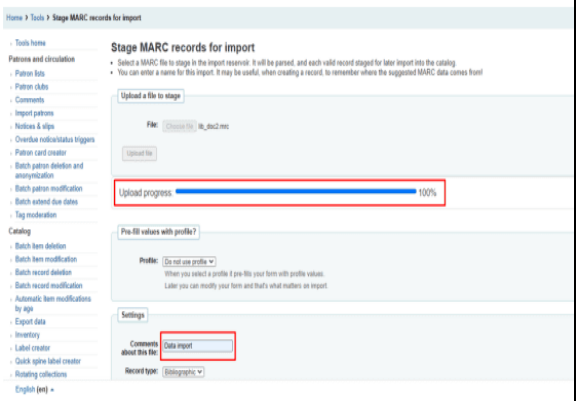

Figure 6d:

After the successful upload of the MARC file to KOHA, a staged MARC record report was provided to facilitate the management of MARC records, making it easier to import bibliographic data, as shown in Figure 6e.

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

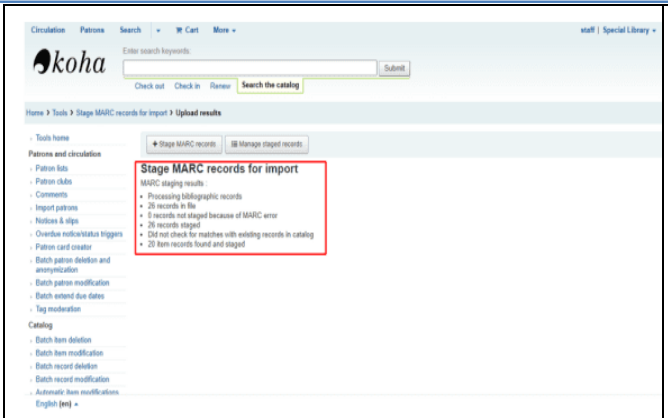

Figure 6e:

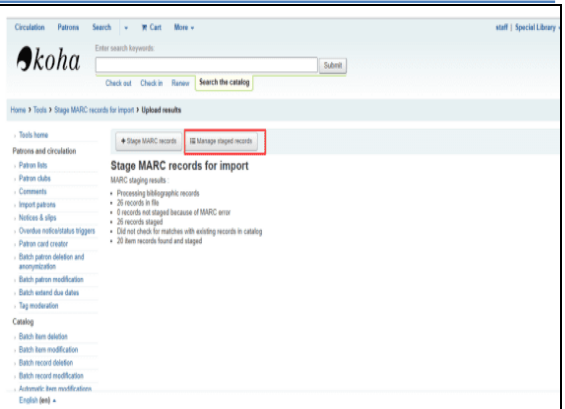

Figure 6f:

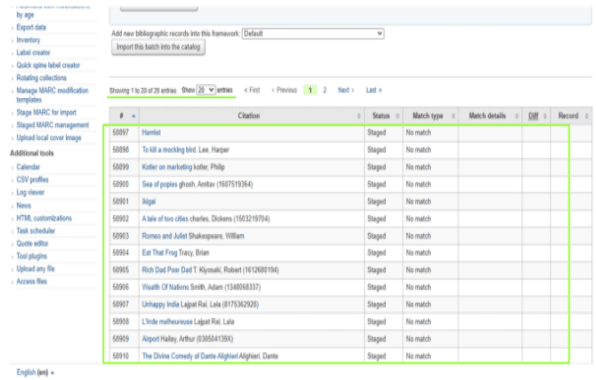

Figure 6g:

The staged records were managed to monitor the flow of data import as shown in Figures 6f and 6g. The data importation process was completed successfully. Therefore, this opened the door for system testing.

Testing of the KOHA system

Immediately after the data migration process was complete, a comparative analysis was conducted to identify errors and missing fields in the KOHA system. The following issues were observed. It was observed that the location of the books could not be migrated fully due to system incompatibility and data structure disparities. Additionally, some of the book copies could not be imported entirely because the data exchange standard used by ADLIB did not conform to the MARC bibliographic framework, which is the standard used by KOHA. Additionally, some of the books migrated into KOHA had incorrect or missing bibliographic records, such as call numbers and barcode numbers. See Figure 7.

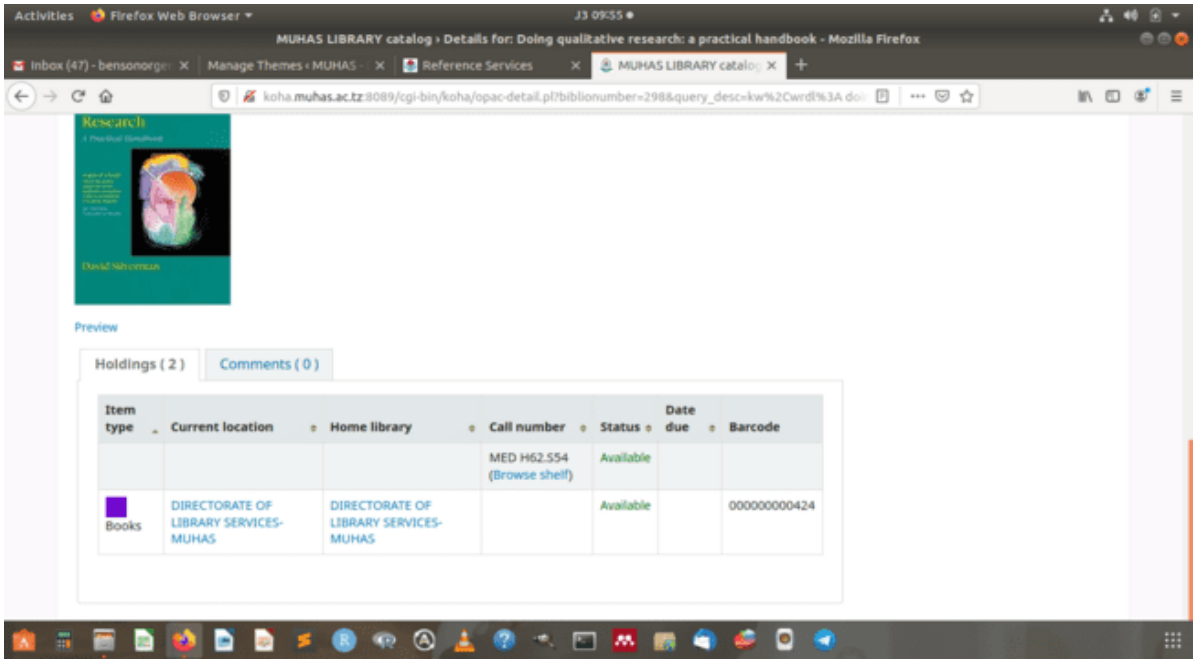

Figure 7: Sample of Book with Missing Call Number and Barcode Number

Results

Success in data migration

The total number of items imported into KOHA was compared to the total number of items in ADLIB to assess the actual success of the migration process. Immediately after the import activity was completed, a comparison was made between the items imported into the KOHA System and the ADLIB System. The total number of items imported successfully into the KOHA System was 22,780 out of 32,313 items that were previously found in the ADLIB System. Therefore, the data migration process was successful in 70.5 per cent of cases. Table 3 presents a comparison between the ADLIB and KOHA systems in terms of selected record information. In addition to the above information, an evaluation was conducted based on specific item details, including title information, Author information, and barcode information.

Table 3: Metadata Comparison between ADLIB and KOHA Systems

S/No	Bibliographic Data	Total number of items in ADLIB	Total number of items in KOHA
1	Authors	32,313	22,780
2	Titles	32,313	22,780
3	Call number	32,313	156
4	ISBN	31,148	26,326
5	Physical distribution	24,942	21,751
6	Date of Publication	27,639	26,295
7	Publisher	21,537	20,641
8	Place of publication	19,913	18,872
9	Dimensions	14,842	13,034
10	Pagination	15,421	12,156
11	Barcode	21,364	234

University of Dar es Salaam Library Journal
Vol 20, No 1 (2025), pp 78 - 100
ISSN: 0856-1818

Challenges in data migration

During the entire bibliographic data migration process performed at MUHAS Library, several challenges occurred and were addressed to reduce migration errors. First, among the challenges encountered was the issue of data mapping and transformation. This aspect presented one of the challenges because ADLIB and KOHA are two different systems with distinct data structures, tags, and formats. Therefore, some item information was missing due to discrepancies in data matching, particularly in terms of data fields and tags. Examples: In the barcode and Call number fields, a very low number of items were imported with both barcodes and call numbers due to differences in data fields.

Second, another data migration shortfall was about the system settings and configurations. This also presented a challenge in data migration, as there were different configurations and settings in some fields. For example, KOHA uses a barcode as the unique Key in its database settings, but this is not the same in ADLIB. Therefore, KOHA did not recognise most of the barcode information from the ADLIB System.

Third, during data migration, we encountered errors related to the size of the data (Data volume) used to import into KOHA. The file size taken from ADLIB to the KOHA system appeared to be of high volume, which led to some activities slowing down and ultimately failing to compile as expected. One of the activities that took a long time was data cleaning, which resulted in unsatisfactory output due to the lengthy process from the Marc Edit software to loading the CSV file.

The fourth issue was about poor initial data entry in the ADLIB system. The initial data entry exercise in ADLIB was performed poorly, as some fields were left blank and others were incorrectly entered. For example, information about pagination was entered into the call numbers data field. The outcome of these anomalies compromised data integrity, hindering the accurate retrieval, organisation, and migration of bibliographic records into the Koha Library Management System.

Discussion

One of the primary causes for the lack of library automation in developing countries is the high cost of proprietary software dominating the market (Hopkinson, 2009; Balaji Babu & Krishnamurthy, 2013). High pricing, a lack of data control, stringent license agreements, and proprietary supporting software all contribute to a higher total cost of ownership for library automation software. However, the presence of open-source integrated library management systems (ILMS) such as KOHA has changed African library automation practices. Nowadays, more libraries are being automated using open-source ILMS, and those using commercial software are opting to migrate their bibliographic data to open-source library software to save the cost of a proprietary license. This is why the MUHAS library did not fall behind in the transition from ADLIB, a proprietary system, to the KOHA library system. Many Tanzanian university libraries have adopted the KOHA library system due to its numerous benefits over proprietary software. Some of the advantages of KOHA include cost savings, ease of use, global community support, increased flexibility, and adaptability, as well as the absence of license fees. This is confirmed by Pandita's (2020) study, which found that libraries ought to shift to KOHA because "*it is not about being free, it is about having complete control*". The primary concern is for libraries to have complete control over their massive data and

information, which means they are no longer dependent on a single vendor or Original Equipment Manufacturer (OEM) for hardware.
The use of KOHA library software has garnered the interest of many university libraries worldwide. For example, over 15,000 academic libraries use the KOHA Integrated Library System (Kulkarni et al, 2023). The system is utilised by public, school, and special libraries worldwide, making it one of the most popular library software solutions. This provides credits to libraries opting to switch from another proprietary system, such as ADLIB, to KOHA, as they are assured of technical support in the event of any operational issues.

To ensure the success of data migration, key considerations must be carefully examined throughout the migration process. Some of the most critical considerations include technical expertise, data complexity and volume, data quality, and tools for data migration. A case study conducted by Todd (2018) demonstrated how academic librarians migrated from Millennium to KOHA, utilising freely available tools such as OpenOffice Calc and MarcEdit. The migration was completed by a dedicated team of librarians who lacked advanced tools and prior experience with data migration or systems integration. As a result of these findings, it appears that having advanced tools is less important than having an expert staff. However, another study found that data migration from legacy software was challenging due to vendor lock-in and non-standard data formats. The findings of this study suggest that the issue of data quality and structure may significantly impact data conversion; therefore, strong teamwork is essential for a successful migration from proprietary software to KOHA.

Data migration can be a daunting and challenging process. Some of the most common challenges include the availability of skilled human resources. Data migration requires a team of experts who are well-versed in the knowledge and skills necessary for the migration process. However, finding the right people with the necessary skills can be difficult. Todd's (2018) study found that, while some libraries rely on institutional IT departments and competent librarians to migrate data, the majority use third-party contractors or vendors. This point of view emphasises how difficult it might be if the migration process lacks the necessary skilled staff. Another factor which may influence the data migration process is compatibility issues between systems. This means that incompatibility between the data formats of the old and new systems can occur when data is moved from one system to another that uses a different database structure or when data is imported into a new system that uses a different file format. The MUHAS library also faced incompatibility in data structure since the ADLIB data field tags were not identical to the KOHA data field tags. Additionally, the ADLIB lacked most of the international standards for data exchange, such as Z39.50, SRU, AACR2, and RDA, among others. This led to the data migration process failing by 29.5 per cent because most of the book copies were not migrated fully. Some of the book's bibliographic details, including call numbers, shelving location, and barcode numbers, were missing.

Moreover, data quality is one of the most critical aspects of data migration. It is essential to ensure that the data being transferred from one system to another is consistent and accurate. Poor data quality can lead to a range of issues, including poor data representation, delays or disruptions in library operational processes, and increased post-migration maintenance costs associated with the tedious work of rectifying errors in the bibliographic database. Data quality issues are exacerbated when migrating from a system with poor data quality to a newer system that offers a set of rich features and a strict data model. This necessitates much planning before the migration process commences. Some of the challenges in data migration projects include ensuring data quality, addressing the lack of data governance, and mitigating

data security risks. To avoid data quality issues, it is crucial to have a data governance plan in place before starting the project. Data governance plans can be defined as a strategic framework that outlines how an organisation manages its data assets to ensure their quality, consistency, security, and compliance throughout their lifecycle. It establishes clear policies, roles, responsibilities, and procedures for creating, storing, accessing, and maintaining data. The data governance plan should outline the roles and responsibilities of the team members involved in the migration process. For the MUHAS library, this issue significantly affected the migration process, as most of the necessary data fields, such as barcodes and call numbers in ADLIB, were incomplete, thereby hampering the quality of data in KOHA. It is essential to note that such crucial details are typically captured under "item information," which falls outside the standard bibliographic framework, thereby complicating the accurate mapping and transfer of data during the data migration process.

Additionally, data migration can be a lengthy and complex process, especially when handling large volumes of data. The time required for migration depends on several factors, including the size of the data, the number of data sources and target systems involved, as well as the speed and reliability of the network connection. One of the most significant challenges in this process is the lack of precise requirements (Hussein, 2021). When expectations are not well-defined, it often leads to misunderstandings and miscommunication between the data migration team and business stakeholders, ultimately affecting the project's success. Furthermore, inadequate data governance can lead to poor data quality and increased security risks (Haug, Zachariassen, & Van Liempd, 2011). To mitigate such issues, it is crucial to establish a robust data governance framework prior to the migration. Overall, a well-defined strategy that includes identifying all relevant data sources, determining the appropriate migration methods, and applying industry best practices can significantly enhance the process's efficiency while minimising disruptions and downtime.

The significance and new knowledge gained from this research can be interpreted in terms of the need for institutional libraries to migrate bibliographic data from one system to another. Migrating bibliographic data between different systems can be a challenging task. It requires careful planning and execution of the planned pathways. As such, this research provided knowledge and insight that contribute to a deeper understanding of how one may develop pathways for an efficient and successful bibliographic data migration process. The core issue here is to reduce manual effort, time, and potential errors during data migration. Findings from this research can therefore be used as a guideline tool for some institutions in Tanzania and beyond that are still using the ADLIB system and are willing to migrate their bibliographic data to KOHA, an open-source Integrated Library Management System. The same procedures highlighted in this study can be applied to most data migration activities from various other systems.

Conclusion and Recommendations

Data migration in library systems is a complex process that requires careful planning. Some challenges affecting the MUHAS library's migration process, transitioning from ADLIB to KOHA, included system incompatibilities, data structure disparities, and poor data quality. Although the MUHAS library was able to migrate 70.5 per cent of the data records, most of the migrated data records were misrepresented, making data modification and rectification extremely time-consuming. Based on the above findings, four interlinked recommendations

are made for consideration when migrating data from one library management system to another successfully.

First, for successful data migration, it is strongly recommended to ensure that the data follows standard cataloguing practices and required formats. Therefore, cleaning up any inconsistencies, duplicate records, and formatting issues is crucial to avoid complications during bibliographic data migration. Second, the activity of matching and mapping data fields between one library management system and another should be well-conducted for the data migration process to be successful. It is also essential to create a mapping document that aligns the fields between the systems, ensuring accurate data importation during the migration process.

Third, it is strongly advised to conduct test migrations with a small subset of data to identify any potential issues or errors before migrating to the entire dataset. Such an undertaking will enable the fine-tuning of the migration process and address any issues earlier. Fourth, a post-migration review must be conducted to monitor the system for any issues and perform necessary system troubleshooting. Therefore, during the post-migration review, it is important to document clearly the issues learned and best practices for future reference.

Acknowledgments

We humbly appreciate the MUHAS library staff and system administrators who participated in this study and provided invaluable contributions.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

The author received no financial support for conducting this research, authorship, and/or publication of this article

References

Ahmed, M. B. (2024). Application of Integrated Library Management Software in Academic Libraries for Information Services Delivery. *Jewel Journal of Librarianship, 10(3): 63-72*

Ahuja, L., Singh, B., & Simon, R. (2024). Data Cleaning: Paving a Way for Accurate and Clean Data. *Global Journal of Enterprise Information System*, *16*(1), 18–25.

Anyaogu, U. (2019). Retrospective conversion of bibliographic records: Koha experience of NIALS library. *Library Philosophy and Practice*, *2019*(December).

Balaji Babu, P., & Krishnamurthy, M. (2013). Library automation to resource discovery: A review of emerging challenges. *The Electronic Library*, *31*(4), 433–451.

Breeding, M. (2018). "Index-Based Discovery Services: Current Market Positions and Trends." Library Technology Reports 54, no. 8 (November/ December): 1–33. https://doi.org/10.5860/ltr.54n8.

COSTECH (2012). Survey Report on the existing IT infrastructure, related environment and

user system requirement for EMIS and E-Library systems, Dar es Salaam.

Das, P.K. (2014). 'Step by Step Guide to Implement KOHA and Process of Data Migration', National Workshop on Library Automation (KOHA), Content Management System (Joomla) & Data Migration, p. 10.

Felicia, U.I. (2015). 'Information needs, library resources and services available to postgraduate students in the Institute of African Studies, University of Ibadan, Nigeria', International Journal of Library and Information Science, 7(4), pp. 77–85. Available at: https://doi.org/10.5897/ijlis2014.0450.

Genoese, L., & Keith, L. (2011). Jumping ship: One health science library's voyage from a proprietary ILS to open source. *Journal of Electronic Resources in Medical Libraries*, *8*(2), 126–133. https://doi.org/10.1080/15424065.2011.576605Haug, A., Zachariassen, F., & Van Liempd, D. (2011). The costs of poor data quality. *Journal of Industrial Engineering and Management (JIEM)*, *4*(2), 168–193.

Hopkinson, A. (2009). Library automation in developing countries: the last 25 years. *Information Development*, *25*(4), 304–312.

Hussein, A.A. (2021). 'Data Migration Need, Strategy, Challenges, Methodology, Categories, Risks, Uses with Cloud Computing, and Improvements in Its Using with Cloud Using Suggested Proposed Model (DMig 1)', Journal of Information Security, 12(01), pp. 79–103. Available at: https://doi.org/10.4236/jis.2021.121004.

Innocent, W. A., & Masue, O. S. (2020). Applicability of E-Learning in Higher Learning Institutions in Tanzania. *International Journal of Education and Development Using Information and Communication Technology*, *16*(2), 242–249.

Kangkum, M. J., & Ogunsola-Saliu, K. (2023). Digitalising the Library Management for Effective Performance in the Nigerian Oil and Gas Sector. *International Journal of Computer*, *43*(1), 73–80.

Karak, S. & Dutta, K. (2017). 'Bibliographic Data Migration in KOHA 3.18 from Existing Data Sheet: A Practical Experience', International Research: Journal of Library & Information Science |, 7(1), pp. 122–132.

Kasulwa, S. (2008). Tanzania Country Report 2008, Conference, Scecsal XVIII Goals, Millennium Development International, Mulungushi Centre, Conference. Retrieved on 12th May 2024 at https://www.scecsal.org/countryreports/tla_report_2008.pdf

Kohn, K., & McCloy, E. (2010). Phased migration to KOHA: Our library's experience. *Journal of Web Librarianship*, *4*(4), 427–434. https://doi.org/10.1080/19322909.2010.485944

Kumar, V. (2015). National Level Seminar on Skills Approach To Information And Communication Technology for Information Problem Solving. St.Mary's College, Thrissur, 2015. Kottayam.

Kumar, V. V., & Majeed, A. K. C. (2019). Data migration from legacy systems to KOHA: A practical approach. *Library Philosophy and Practice*, *2019*, 1–17.

Luena, A. (2012). 'Strengthening the Education Management Information System (EMIS) in Tanzania: Government Actors' Perceptions about Enhancing Local Capacity for Information-based Policy Reforms', Master's Capstone Projects, p. 58. Available at: https://scholarworks.umass.edu/cie_capstones/21.

Machovec, G. (2014). Consortia and Next Generation Integrated Library Systems. *Journal of Library Administration*, *54*(5), 435–443. https://doi.org/10.1080/01930826.2014.946789

Malik, K.M. (1996). 'The status of library automation in Pakistan', Library Review, 45(6), pp. 36–42. Available at: https://doi.org/10.1108/00242539610125668.

Manda, P. (2003). 'Planning for automation of the University of Dar es Salaam Library', University of Dar es Salaam Library Journal, 5(2), pp. 1–15.

Masenya, T. M., & Ngulube, P. (2021). Digital preservation systems and technologies in South African academic libraries. *South African Journal of Information Management*, *23*(1), 1-11.

Muneja, P.. (2010) 'A reflection of Tanzanian libraries in the digital age: Challenges and prospects', *A Paper Presented as a Key Note Address During the Annual Conference for the Tanzania Library Association (TLA) held in February, 2010 in Mbeya Tanzania.*, (October), pp. 6–9.

Mushi, C., Mwantimwa, K. & Wema, E. (2020). 'Perceptions of librarians towards the adoption of embedded librarianship in Tanzania', 15(2), pp. 21–38.

Otike, F., & Barát, Á. H. (2021). Roles and emerging trends of academic libraries in Kenya. *Library Hi Tech News*, *38*(7), 19-23.

Pankaja, N., & Mukund Raj, P. K. (2013). Proprietary software versus open source software for education. *American Journal of Engineering Research*, *2*(7), 124-130.

Potter, C. (2007). 'This survey shows the market for data migration is very significant … and cost and time eliminated', Europe [Preprint], (September).

Reddy, T. R., & Kumar, K. (2013). Open source software's and their impact on library and information centre: An overview. *International Journal of Library and Information Science*, *5*(4), 90–96.

Sahoo, S. & Saikia, M. (2019). 'Data Migration from Libsys to KOHA ILS: An Experimental Study', SRELS Journal of Information Management, 56(April), pp. 85–90. Available at: https://doi.org/10.17821/srels/2019/v56i2/141234.

Samzugi, A. S. (2016). Status of library automation in Tanzania's public universities. *University of Dar Es Salaam Library Journal*, *11*(1), 24–38.

Tella, A., & Oladeji, T. I. (2017). Empirical investigation on the impact of KOHA on library services in selected academic libraries in Nigeria. *Annals of Library and Information Studies (ALIS)*, *64*(2), 113–115.

Thalheim, B. & Wang, Q. (2013). 'Data migration: A theoretical perspective', Data and Knowledge Engineering, 87, pp. 260–278. Available at: https://doi.org/10.1016/j.datak.2012.12.003.

Todd, C. R. (2018). Librarian as data migrator: a functional pathway from Millennium to KOHA. *Digital Library Perspectives*, *34*(1), 60–69. https://doi.org/10.1108/DLP-09-2017-0035

Vimal Kumar, V. & Abdul Majeed, K.C. (2019). 'Data migration from legacy systems to KOHA: A practical approach', Library Philosophy and Practice, 2019.

Vimal Kumar, V., & Majeed, K. C. A. (2019). Migration from SOUL to KOHA: A Learning Experience. IndKoha 2019, February, 33–37. http://eprints.rclis.org/34019/1/Migration from SOUL to KOHA a learning experience.pdf

Wamunza, N.A.G. (2003). 'Library Automation: The Role and Significance of Library Automation Plans', *Huria Journal of the Open University of Tanzania*, 5.