
Some Tech Know-How Needed

Skill Building for Migration to a Hosted Open Access System at an Academic Library

by Jennifer Ye Moon-Chung, Amy Lee Heinlen, Kristy White and Rob Behary

INTRODUCTION

With the advent of the software as a service (SaaS) and platform as a service (PaaS) models for managing integrated library systems (ILS) comes a reduction in responsibility for individual libraries to solve many technology-related complexities. In SaaS/PaaS models, the onus of implementation and ongoing management of the ILS falls to employees of the selected software and platform provider; library staff are no longer tasked with challenges like acquiring necessary hardware, creating an appropriate cloud-based environment and installing the appropriate operating system/web server.

This advancement in service models results in significantly less expertise and training requirements for library employees. Libraries no longer need employees to manage hardware, patch software, and apply security updates. However, despite the benefits of an ILS hosted in a SaaS/PaaS environment, significant technology skills are still required to fully maximize its potential for libraries. These skills prove necessary to ensure efficient loading and exporting of data, interaction with the institution's enterprise resource planning (ERP) software, and statistical retrieval and manipulation for data produced by the ILS.

The following article highlights some of the skills librarians need to successfully implement and maintain an ILS in modern library systems architecture. Platform as a service (PaaS) includes the systems architecture on which the library software will run, and software as a service (SaaS) refers to the library ILS software itself. The authors have divided the skills needed to successfully interact with both the PaaS and SaaS architecture as follows: importing and exporting data, interacting with enterprise/external systems, and generating reports. Finally, the authors share pathways for acquiring the skills needed to accomplish these tasks.

SKILLS NEEDED FOR IMPORTING AND EXPORTING DATA

A primary component of any migration is to make sure that the data can be extracted from one system and loaded into another system without introducing any errors. To some degree, the SaaS/PaaS service providers can be of assistance, especially regarding the process of loading the data, but for data extraction purposes, librarians must often rely on their own knowledge and expertise to extract the data from their existing system. The skills that may be most helpful to build for data extraction are query languages such as SQL and basic programming skills in a language such as Python.

Jennifer Ye Moon-Chung, M.Ed., is the Digital Projects & Analytics Manager at Duquesne University's Gumberg Library. Amy Lee Heinlen, MLIS, MFA, is Assistant University Librarian for Access Services at Duquesne University's Gumberg Library. Kristy White, MSLS, MS, is Head of Licensing & Acquisitions at Duquesne University's Gumberg Library. Rob Behary, MLS, MBA, is Head of Systems & Scholarly Communications at Duquesne University's Gumberg Library.

SQL is Structured Query Language, which is meant to be interoperable with many database systems. In some cases, vendors may provide access and descriptions of tables that can be used to decipher where the data is located within the database tables of the ILS, and vendors will provide the ability to connect to the ILS with a standard SQL retrieval tool, but there is still a learning curve associated with SQL. Novice users may miss important aspects of the table relationships where data cannot be extracted with a very simple query. Fortunately, courses in SQL are offered at many universities. There is also a wealth of information available about SQL through the various websites where the variations of SQL reside.

Similarly, knowing a programming language such as Python can be very helpful in the migration process. Python, and especially its Pandas module, can be used to manipulate migration data without introducing some of the errors that are commonly introduced by programs such as Microsoft Excel, where rounding errors and truncation issues can occur during the editing process. Python and Pandas do not introduce such errors, and the environment required to run Python is relatively simple to manage or learn. It is common for Python to be taught at most universities, and Python has become an essential component of many library programs.

SKILLS NEEDED FOR INTERACTING WITH ENTERPRISE/EXTERNAL SYSTEMS

Enterprise systems at universities contain student records, accounting transactions, and other types of data needed by the ILS. External systems also include ordering software from library vendors, software from library consortiums that allow libraries to exchange materials, and protocols that can harvest the records of one system to be included in another system. Librarians may not need to master these technologies completely, but a working knowledge of each of them can be most useful in troubleshooting problems and ensuring accurate interchange of data with enterprise and external systems.

The Security Assertion Markup Language (SAML) and the Lightweight Directory Access Protocol (LDAP) are two methods by which the identities of library users are exchanged and verified at many universities. SAML especially has become a standard as more and more universities become members of federated identity programs through Internet2 memberships. Both SAML and LDAP transport personal information in a standard way so that the information can be used by the library to provide access to account information and validate library users of electronic material. SAML is XML-based and requires some knowledge of how records containing personal information are structured by a university's enterprise system. The LDAP protocol, which expects records to be structured according to the LDAP Data Interchange Format (LDIF), requires special field names and mappings similar to SAML. Librarians should have at least enough knowledge about both SAML and LDAP to be able to troubleshoot problems with the transfer of personal information, but a deeper knowledge of SAML and LDAP can lead to creative delivery of services and potentially new opportunities to improve the functionality of the ILS.

Application Programming Interfaces (API) and JavaScript Object Notation (JSON) are other technologies that allow for interaction with external systems. When paired with JSON, APIs allow for the exchange of financial information, and increasingly personal information, between enterprise and external systems and the ILS. APIs can work with most major computing languages, including Python. While APIs provide the method for exchanging data in these situations, JSON is the structured mechanism that encodes the data itself. A knowledge of both can open a world of possibilities that allow for the easy exchange of data with enterprise and external systems.

SKILLS NEEDED FOR REPORTING

As we have discussed, migrating to a new library system involves various technical skills, but it is equally important to possess the necessary skills for effective reporting. Without the necessary technical skills, such as programming skills, librarians may struggle to integrate data from different sources and clean it to increase accuracy, leading to a lack of clarity and valuable insights in their report.

Programming skills like Python and R are particularly useful when handling different data types and formats. Customized Python code can be created to convert data structures, extract necessary information, and match data to new fields, while also removing sensitive information. Utilizing libraries such as Pandas and NumPy enables librarians to easily remove duplicates, handle missing values, standardize data, and reorganize datasets. This data cleaning process minimizes errors and helps to derive more meaningful summary statistics.

Similarly, R provides several library options to manipulate, reshape and transform data, such as dplyr and tidyr. Additionally, R provides easy-to-use visualization libraries, like ggplot2 and shiny. These allow for creating interactive web applications along with the integrated software, RStudio. It is a great idea to master the entire process from data preparation to data visualization using the tools provided by R.

In addition, PowerBI and Tableau can be great options to visualize data and answer meaningful questions. Libraries can use these tools to visualize their data in a way that is easily understandable and actionable. For instance, a library might use one of these tools to create a dashboard that shows circulation statistics over time, broken down by patron demographics or types of materials borrowed. In these ways, librarians can be prepared to generate the reports from the PaaS/SaaS environment in any way that the vendor allows.

PATHWAYS TO ACQUIRING SKILLS

With the continuous advancement of technology, it is imperative for librarians to continually update their skill sets. This requires a commitment to ongoing learning and development of new skills, as well as a willingness to adapt to changing systems and processes. While there are various pathways to acquiring these skills, choosing the best way to learn depends on one's circumstances and learning preferences.

Formal education is an excellent option for those seeking to acquire technology skills. Graduate programs in Information and Library Science or Data Science, as well as certifications in Data Analytics or Programming, are examples of formal education pathways that can be especially beneficial for new librarians.

User communities are another effective way to learn new skills. These communities offer opportunities to learn from peers and to network and collaborate. For instance, the FOLIO user community is a big network of librarians, developers, and service providers working together to develop an open-source library services platform. Members can access a variety of resources, such as forums, webinars, and workshops, to learn from one another and share their expertise.

Additionally, informal learning opportunities, such as online tutorials and courses, and self-directed learning through books or blogs, are great ways to stay current with technology skills. These resources offer flexibility and the ability to learn at one's own pace.

Many PaaS/SaaS vendors will also offer formal learning opportunities to acquire skills use in a particular PaaS/SaaS environment. These types of classes may be for the API that interacts with a particular program, an explanation of the database tables that describe how data is stored in a particular system, or an outline of the structure of how to load files into a system. While sometimes there is an added expense for these types of classes, the benefits of knowing the idiosyncrasies of a particular system can be beneficial in the long run.

CONCLUSION

In conclusion, while the adoption of SaaS/PaaS models for managing ILS has relieved libraries of hardware and software management responsibilities, librarians still need to acquire and maintain technical skills to effectively interact with the new technology architecture. They need to develop skills in importing and exporting data, interacting with enterprise/external systems, and generating reports. Familiarity with SQL, Python, SAML, and LDAP can be particularly helpful in these areas. By acquiring these skills, librarians can ensure efficient data transfer and retrieval, troubleshoot problems, and potentially even improve the functionality of the ILS.