

Current and Future Library Catalogs: An Introduction to FOLIO

by Peter McCracken (Electronic Resources Librarian, Cornell University) <phm64@cornell.edu>

Over the past ten years, many academic libraries have implemented new online catalogs and automation tools, to better manage the significant shift in materials expenditures from print to electronic. At **Cornell University**, where I work, the University Library spent nearly 70% of its 2016/17 materials budget on electronic resources. Many institutions spend a much higher percentage on electronic resources. But too many libraries — including **Cornell** — do not yet have tools that can effectively manage the result of this significant shift in spending. The online catalogs of previous generations are not well designed to manage this change, and in most cases libraries have found it necessary to use one or more separate tools to manage all of this spending. Too many tools, and too much time, is spent dealing with checking and correcting links, managing licensing details, trying to determine actual holdings rights, ensuring accurate access for the correct individuals, and much more.

In the past decade, several new library automation systems have arrived on the market to better manage this shift. The most common are **Alma**, from **Ex Libris** (now owned by **ProQuest**); **Sierra**, from **Innovative Interfaces**; and **Worldshare Management Services**, from **OCLC**. These new systems have seen significant adoption; as described below, among 205 leading college and university academic libraries in the United States and Canada, 122, or 60%, have shifted to a new system in the past ten years, and 97% of those installs (118 of 122) were one of these three systems.

This shift will continue in the next few years, and over the next year, **FOLIO** — a new, open-source library management service — will become available to libraries, as well. My employer, **Cornell University Library**, is fully committed to **FOLIO** and, along with several other libraries, has committed significant resources to its development. **FOLIO** will introduce the next evolution to the library

management service marketplace, and one very much worth libraries following and considering.

Since the introduction of the first online public access catalogs (OPACs) in the 1980s, through the integrated library systems (ILSs) of the 1990s and 2000s, and to the library management services platforms (LMSs) of today, libraries have sought to find effective tools for making their collections accessible to their patrons. It has not been easy, and each transition has tried to improve upon the problems of prior systems. The domination of electronic resources in the library marketplace has meant a change in how librarians offer, and manage these resources. The proliferation of electronic content means libraries generally cannot keep track of exactly what is available in the databases to which they have access. They must rely on outside companies to track this information, and tracking that information is now a critical part of the library management system. Managing licensing — and determining who should be able to access what, via which channels — has become a vital part of the librarian's skill set, and we need tools to help us manage that work. In addition, allowing appropriate access without leaking too much personally identifiable information, becomes vital. All of these, and many other functions, require a completely new toolset.

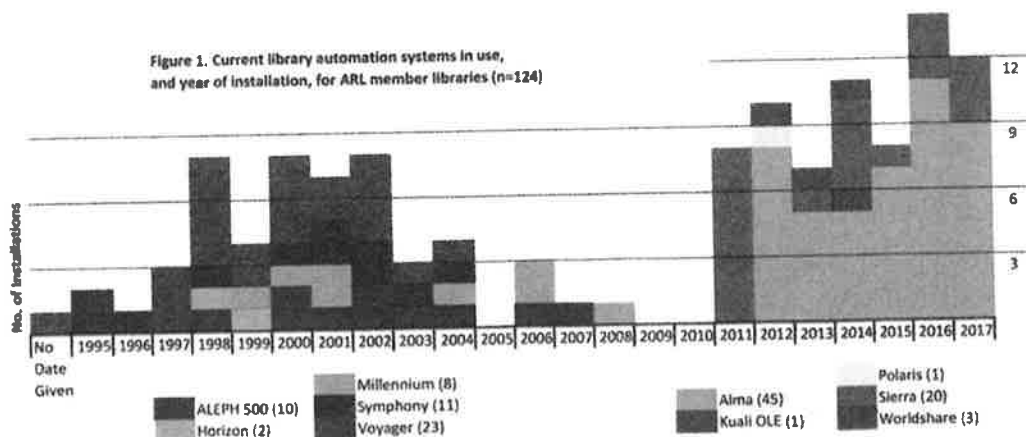
In thinking about upcoming changes in the library automation universe, I wondered who is using which library automation systems, and how long have they had their current system — or, more specifically, how many libraries have *not* upgraded to a better system? Using data collected from **Marshall Breeding's** valuable database about library automation tools, at *LibraryTechnology.org*, in late January 2018, I reviewed the current systems in use in Association of Research Library members and Oberlin Group members. I sought a way to display the age and system

in use at those institutions; hopefully, the attached figures will be useful in doing so.

Figure 1 shows a summary of all current automation systems among members of the **Association of Research Libraries (ARL)**, representing the largest research libraries in the United States and Canada. A clear split exists between the prior generation of automation systems, generally known as ILSs, and the current generation, or LMSs. Among 125 ARLs, one does not currently have an automation system. Of the remaining 124, 70 (or 56%) are using catalogs acquired in the last eight years. The division between the two generations of catalog systems is undeniable; no system appears on both sides of the 2009/2010 gap, when no new systems were installed. The vast majority of these new installations, among ARLs, was **Alma** or **Sierra**. Note, however, that 40% of the **Sierra** installs were in 2011 alone; in contrast, **Alma** has had a regular schedule of at least 5, and an average of 7½, ARL installs, per year, since 2012. **Alma**, and most likely its electronic resources management features, have appealed to many research libraries.

While close to 60% of all ARLs acquired a new automation system in the last eight years, over 40% did not. Of those, nearly half (23 of 54) are using **Voyager**, from **ProQuest**. All of these libraries are almost certainly looking for new catalogs that will allow them to do much more, more effectively, than they currently can. While electronic resources obviously existed when these automation systems were installed, the world of resource acquisitions has changed dramatically, and in many research libraries the electronic resources budget has grown from perhaps 20% of all expenditures to some 80% of all expenditures. The tools that libraries need to manage these electronic resources simply do not exist in the catalogs they use, and most libraries, if they have some sort of electronic resources management tool, must use that tool outside their ILS. Some libraries — for example, nearly all university libraries in Germany — still do this work in basic spreadsheets and home-made databases. Other libraries use tools clearly not intended for this work, such as **Trello**, the online project management tool. While libraries are able to make parts of it work, these tools have many gaps that libraries need filled. Still others, such as **Cornell University**, are using **Intota**, an electronic resources management tool from the **Serials Solutions** division of **ProQuest**, but there are no plans to grow or support **Intota**, as **ProQuest** not sur-

Figure 1. Current library automation systems in use, and year of installation, for ARL member libraries (n=124)



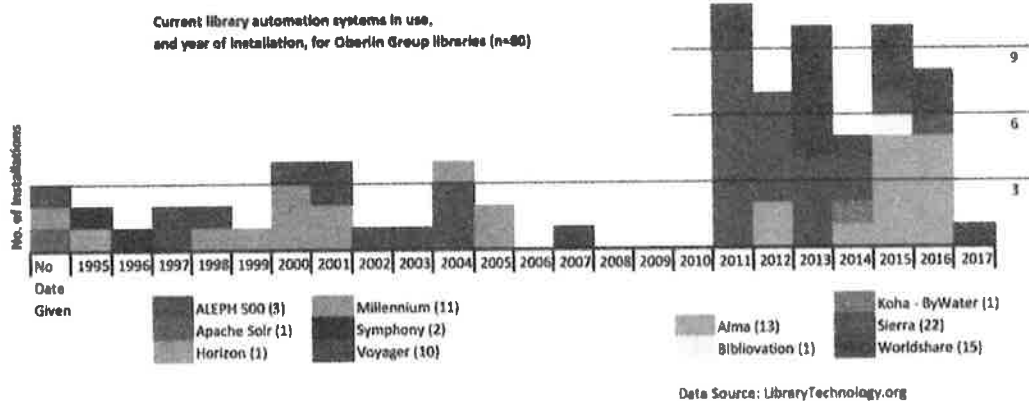
Data Source: LibraryTechnology.org

continued on page 14

prisingly hopes to migrate these libraries to Alma, instead.

The situation is similar in smaller American academic college libraries. As shown in Figure 2, the Oberlin Group of libraries had an even more noticeable break between the previous automation systems of pre-2008, and the newer systems since 2011. The Oberlin Group is a gathering of 80 selective college and small university libraries. The group serves as a good indicator of the direction of college library automation plans.

Between the start of 2006 and the end of 2010, only one Oberlin Group library installed a new automation system. (It is possible that a library acquired one in that period and then replaced it between then and now; such a system would not be included in this chart.) Obviously, some of this gap resulted from the Great Recession of 2008 and its aftermath, though this period includes two years before the downturn began in December 2007. Once an effective solution existed and their financial situations allowed them to acquire a new system, many college libraries did so. In almost every case, they selected either **Innovative Interface's** Sierra (in 22 cases), **OCLC's** Worldshare Management Service (15), or **Ex Libris/ProQuest's** Alma (13).



But as with major research libraries, many small college libraries still need to implement a better, more effective, solution. Those libraries that are still using past-generation catalogs need a collection of tools that will ensure they are providing access in the most efficient and most effective manner possible. Many are certainly considering Alma or Sierra (or, for the smaller institutions, often Worldshare Management Services), but some are also awaiting the completion of FOLIO, the open-source LMS being developed by the **Open Library Foundation**, with extensive financial support from **EBSCO**, the **Mellon Foundation**, participating libraries, and many other funders.

FOLIO represents the next evolution in the field of LMSs, and is being developed by a wide range of libraries and vendors from around the world. FOLIO is open-source, so when released it will be freely available to

all. Expanding on ideas and tools from the past decade, but being built brand-new from the ground up, FOLIO provides a new way of managing library access to all resources. FOLIO is an expansion of the Kuali OLE project to build an open-source library management system, and financial support from **EBSCO**, along with other partners, has created a community of developers committed to creating a management system that any library can use.

The end product will be free and open-source, in the sense that anyone can download the software to run the system. But as is often said about open-source projects, "It's 'free,' as in a free puppy; not 'free' like free beer." To make FOLIO work, libraries will still need to commit time and resources to install, support, implement, and maintain the system. **EBSCO** and others will offer those support services to the libraries that seek it, for a fee.

From a business point of view, this highlights one of the biggest differences between FOLIO and existing LMSs. For stand-alone systems like Alma or Sierra, libraries can only acquire the product from the producer, and while there may be some price competition between different products, there is no competition for installations of, say, Alma, since **ProQuest** is the only provider. Similarly, long-term support for, and access to, the resource can only come from the original provider. Long-term contracts are critical for protecting a library from unexpected price in-

creases. For an open source product, anyone can support the product. This could include a library vendor like **EBSCO**, **SirsiDynix**, or **Bywater Solutions**, which has been supporting the open source Koha system for years. Other viable options include local IT firms, or a single institution on its own, or perhaps an IT team from a consortium, which could support all of the consortium members' installations. Given the many options for installing and supporting an open source solution, one can reasonably expect a significantly lower annual maintenance cost than one would see from a single-source provider.

FOLIO is being built by programmers, developers, product owners, product managers, subject matter experts, and many more, from a range of institutions in North America, Europe, and Asia. Many developers work for **EBSCO** or companies hired by **EBSCO**

to build various parts of FOLIO. Still others work for participating libraries — **Cornell**, **Duke University**, and **Texas A&M University**, in particular, have developers dedicated to writing FOLIO code, while many, many, others have subject matter experts who contribute time and expertise through Special Interest Groups that define how the product will be built. Funding from the **Mellon Foundation** and others also provides support for additional developers and product managers through the **Open Library Foundation**, which oversees and directs FOLIO development.

The FOLIO project aims to have an initial "minimum viable product" available for use by July 2018. By January 2019, many additional units will be complete, with much more functionality added. A full "version 1" should be available for download and use by next January. Of course, any major project like this takes an enormous amount of work to complete, and the first versions are intended for libraries that are willing to contribute beta testing time and experience. But with that comes the ability to directly affect the manner in which the project develops; in the end, those who participate from the early stages will have the opportunity to directly impact how the product grows and improves.

Two parts of FOLIO that will be of particular interest, I believe, will be the **Codex** and the **FOLIO Marketplace**. The **Codex** is among the most difficult to understand, and I must admit that the description that follows comes from time spent discussing the **Codex** and its structure with a collection of expert developers — specifically, the ones creating the **Codex** — but even after all of that, my understanding may be limited or inaccurate, or the eventual structure and actual implementation of what is expected of the **Codex** may change.

As I view it, **Codex** is a concept that represents the ability for a FOLIO user (generally, a library employee, not a patron) to search a vast world of resources through a single interface. Separately, however, is the ability to search a subset of **Codex** data — generally, but not always, referred to as an institution's "Inventory" by FOLIO developers — that represents all of the resources that an institution chooses to make available to its users or patrons. This **Inventory** will be compiled within FOLIO, and will be the content behind most of the activity relevant for a library's daily work, such as circulation, licensing, access, ILL activities, cataloging, inventory, holdings management, and more. At present, however, there are no plans to create a public interface to this FOLIO **Inventory**, so libraries will generally use a tool such as **VuFind**, **Blacklight**, or perhaps a commercial discovery layer, to provide patron access to the institution's **Inventory**.

continued on page 16

Each institution defines what is in-scope for its Inventory. This would almost certainly include descriptive records for the physical volumes that the library owns and stores on its shelves, along with electronic records for the resources that a library subscribes to, purchases, or leases, for and on behalf of its patrons. At the institution's discretion, the Inventory can also include many other subsets, such as descriptions of — or perhaps the full text of — institutional repository data, electronic theses and dissertations, discovery layer contents, resources that are available for patron-driven acquisition, externally hosted eBook and audiobook collections, institution-specific online content from commercial publishers, and much more.

A library might provide different collections to different individuals, through the Codex search tool. For example, acquisitions staff might have access to data stores far beyond what appears in the patrons' Inventory, so they can quickly and easily determine sources for the acquisition of print and electronic resources. An acquisitions staffperson might use the entire FOLIO Codex tool to identify all the vendors who could provide an electronic version of a particular monograph, and may also discover that the monograph is already available through a path that perhaps the patron or public services staff had not located. But the Codex is limited to the resources that are willing to be included in it; if a vendor of eBooks chooses not to allow its data to be searched via the Codex, then its results will not appear there, and the acquisitions staffperson will need to search that resource separately if they want to consider its contents.

Each collection could essentially be a separate knowledgebase that the Codex can search, if or when a library chooses to include it. Or, the library might decide to track all of its electronic resources in a single knowledgebase. Because the FOLIO project is creating and defining specific APIs about how a knowledgebase is searched, it is up to each knowledgebase vendor to offer access to their resource through the FOLIO structure. Currently, EBSCO has built an eHoldings app that will allow librarians to manage electronic resources data in their own EBSCO knowledgebase tool. Other knowledgebases, of nearly any size, can create interfaces between their data and the FOLIO system, using the appropriate APIs. Or, a third party could build and distribute a tool that allows a library to access a vendor's data through a combination of that vendor's data feed and the third-party's app or interface. Librarians will have many opportunities to access and manage the data that most matters to them.

Generally, the Codex search of all available resources does not search MARC records, per

se. "Codex" searching will be limited to a small number of (as yet, not-completely-defined) fields, such as author or contributor, title, some relevant publication data, and perhaps some subject data, when available. It will, in essence, be nearly like searching BIBFRAME "Instances" — looking at the concept of a resource, but not necessarily the "Manifestation" level common to the FRBR structure. But Codex searching *will* link out to more descriptive item-level data from other resources. So a search for a specific monographic title will locate a brief record that represents the title, and then will have perma-links to knowledgebases that provide more information about that resource, such as representations of electronic versions of the title, or print copies in the library's physical inventory, or consortial copies available from other affiliated collections, or information about copies that can be purchased or leased — all depending on which collections or volumes or knowledgebases the library chooses to include for its users.

When looking at bibliographic data, this information will, invariably, be MARC records — but only because that's what is so prevalent today. However, the Codex structure does not specifically demand a MARC format, and any other appropriately-structured data could be included in an institution's Inventory.

If, for example, a library chose to start using Resource Description Framework (RDF), or some other encoding standard, to describe the resources it offers to patrons, all newly added content could be described with RDF, and existing content could

remain in MARC. The Codex would search both collections, and results would appear to patrons as just a single Inventory. Over time, MARC records could be transformed to the newer encoding standard, and the data would shift from one collection to the other — though the shift would not be obvious to the end user.

Like the Codex, the Marketplace will be a new space that brings together disparate resources into a single environment. But instead of bibliographic data, the marketplace will offer tools, workflows, services, data, and other functionality in an environment that will provide many parties with the opportunity to buy, sell, offer, trade, or implement these tools as they see fit. The Marketplace's best comparison is almost certainly to an App Store, in that the resources that appear in the Marketplace are designed specifically for the FOLIO community, and contributors can decide if, or how much, they will charge for the products and services they offer in the Marketplace. User reviews can help guide others toward or away from specific tools that might or might not meet their needs.

Many libraries may prefer to stick with an automation tool from a single-source vendor; certainly, the creator of the product most likely employs the most knowledgeable individuals regarding the resource in question. Others, though, may prefer the flexibility that an open source solution provides, both in terms

of finding companies to support the library's installation, and in finding additional tools to support the product in ways that the single-source vendor may choose to not do, or may not be able to do.

Because it is open source, libraries will be able to use just the modules that interest them, and if they feel that a certain unit does not meet their needs, they can select a different module with a similar function, or even build their own. If the tools that currently exist don't meet their needs, they do not need to wait until the vendor decides to get around to building the tools they need — they can build it themselves, or hire someone to do so, to their specifications. They could, then, offer their solution to others, charging money for it or not, as they see fit. While the parts of FOLIO that are currently being built — by OLE, EBSCO, and many smaller companies — are "open source," so anyone can download and use them, this does not mean that all aspects of tools for FOLIO must be either open source or free. Within the FOLIO Marketplace, companies, individuals, libraries, and others, can all build and sell or give away tools that they believe will benefit the library community.

Similarly, within the standard FOLIO tools, people, libraries, and companies will be able to build and sell or distribute workflows that will help libraries manage their resources. For example, a library that wants to track all arrivals of a print serial with a complex publication schedule could purchase a workflow that was designed specifically to track that particular serial. Or someone could build and share the steps necessary to correctly establish access to specific complicated electronic resources. Another vendor could create and offer a tool that helps libraries identify missing online resources, or perhaps a tool that checks availability of online resources. Libraries need many, many tools to manage everything they offer to patrons, and the FOLIO Marketplace will make it easier for smaller vendors to get their useful tools to the libraries that need them. In every case, it will be up to the creator to decide the cost and availability of such workflows and other tools and apps.

The library automation marketplace continues to evolve. Libraries expect more efficient and more effective ways of managing the resources they offer to their patrons, and it should be no surprise that managing electronic resources is among the most important services needed today. Past-generation systems simply do not offer this functionality, and libraries find themselves working hard to force old or inappropriate tools to do this difficult work. The current generation of systems, particularly Alma, Worldshare Management Services, and Sierra, provide librarians with the tools they need.

But the next generation tool, FOLIO, will provide new and additional functionality to manage these resources in a more efficient manner. And, more importantly, FOLIO will introduce a broad paradigm shift in how libraries run these systems. As an open source

continued on page 18



Optimizing an Octopus: A Look at the Current State of Electronic Resources Management and New Developments in the CORAL ERM System

by Heather Wilson (Acquisitions & Electronic Resources Librarian, Caltech) <hwilson@caltech.edu>

No one mentions in job descriptions that being an electronic resources librarian means drawing a lot of octopuses. In addition to dabbling in supply chain management, scholarly communication, and acquisitions, e-resources professionals are often asked to explain why those areas are related to the work of managing electronic access. The result is often a series of sprawling, tangled diagrams with arms reaching everywhere, all in an attempt to represent the “electronic resources management workflow.” The unfamiliar viewer may be shocked by the complexity of processes; knowledgeable viewers are often only marginally less overwhelmed. The reality is that electronic resources management (ERM) is dynamic, unstable, and unpredictable, and that is if the librarian is lucky enough to work in a library that promotes innovative approaches to those processes.

Librarians and library software companies have dedicated considerable resources to developing systems that support and simplify the library’s ERM workflows, especially as the systems have multiplied beyond the central ILS. In addition to holdings and licensing information, many libraries have added the link resolver, the proxy server, web scale discovery systems, helpdesk ticketing systems, and other tools to their suite of services, all of which have a role in the e-resources life cycle. However, most attempts to simplify and centralize these individual services have been unable to encompass emerging tools and ultimately generate the need for more extended workflows, more octopus diagrams, and more spreadsheets in a shared drive.¹

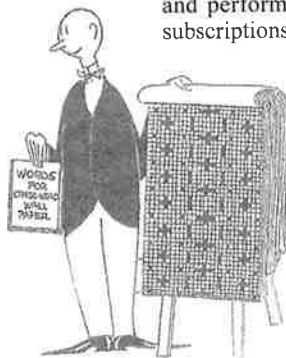
It’s been nearly a year since **Emily Singley** and **Jane Natches**, two researchers from **Boston College**, published

about the pervasiveness of those external ERM processes at institutions with Library Service Platforms (LSP) in place, research which illuminated some of the central challenges in developing ERM systems.² LSPs, also commonly called “Next-Gen ILSes,” have emerged as one approach to managing both electronic resources and library systems, an approach which hopes to reduce the number of processes by centralizing them in one larger, often shared system. Libraries have seen “a growing trend toward a consolidation of services for electronic resources management, A-Z journal listings, full text link resolving, and discovery services under a single service provider,” and these centralized platforms are a natural evolution of that trend.³ Using the TERMS framework (<https://library.hud.ac.uk/blogs/terms/>) as a guide for some of the more universal functions of ERM, **Singley** and **Natches** conducted a survey of nearly 300 library professionals to uncover how ERM was being handled in three major LSPs: Alma, Sierra, and OCLC WorldShare. The survey respondents expressed that the systems have simplified many workflows, especially with regard to journal activation within the central systems. Nevertheless, all of the surveyed library staff expressed the need for performing many ERM tasks outside of the systems, particularly in managing renewals and performing ongoing assessment of the subscriptions. Across all three systems, more than half of staff users were assessing their renewals outside of the LSP in spreadsheets, shared drives, and other external options.⁴ Usage modules and cost-per-use calculations continue to be a struggle to manage within central systems, and even many tools designed specifically for this purpose are not able to provide the full range of COUNTER data and

calculations. For example, **EBSCO’s** Usage Consolidation tools provide excellent ways to centralize JRI, BR1, BR2, DB1 and DB2 reports;⁵ however, many libraries are finding value primarily in the JR5 and JRI GOA reports.^{6,7} This is not to single out that particular tool; finding systems to handle the calculations cited above is challenging, and demand for these reports and calculations is only just emerging. While improving these tools is a constant process, meeting these increasingly more specific needs at the current pace of collection management is a challenge for a centralized system, where other workflows will inevitably be affected.

The related challenge of feasible data migration in a major system change continues to be a discussion point in the centralization of these processes. Although librarians may appreciate the minimized processes that come with one-stop ERM in an LSP, that value is often contingent on how carefully the e-resources and metadata functions are being handled. **Singley** and **Natches** note that the “complexity and ever-changing nature of ERM has made it necessary for libraries to invest in multiple software systems as well as use manual workarounds to support ERM workflows.”⁸ Many of the systems were not originally designed for interoperability, particularly in the parsing of metadata between them. This lack of connection between systems has often resulted in siloed workflows that are difficult to centralize without extensive project management and staff time. Libraries in this and similar situations have begun to explore a different solution than system centralization: perhaps what is actually needed are configurable systems and integrations that allow ERM to be implemented more gradually, developed along the library’s own timeline. This solution looks instead at creating connections between existing systems through APIs and integration tools, which will allow a library to make system changes only when necessary (and reasonable to do so).

The details of the open source FOLIO LSP had not yet been announced when **Singley** and **Natches** performed their survey in early 2016, but presentations of FOLIO at the recent 2017 **EBSCO** User Group reflected library demand for stronger integration of pre-existing platforms and clearer data migration plans.⁹ In fact, the agendas from the user groups of other major LSP providers, including **Ex Libris** and **OCLC**, included sessions about API development and fluid integration with outside systems.¹⁰⁻¹¹ It’s clear that many librarians and library providers are coming to the conclusion, as **Singley** and **Natches** did, “that ERM remains a complex process that is, as yet, too daunting to encompass within any



Current and Future Library ... from page 16

product in a community that values new ideas and new tools, FOLIO will, I believe, provide the opportunity for libraries and small vendors to develop and offer tools that will benefit the entire library community — and its patrons — in ways that we cannot begin to imagine.

As we finally find and develop more and more useful ways of ensuring that our patrons will be able to truly take advantage of all the

resources we acquire and offer to them, this is an exciting time to play a role in defining the next generation of these critical tools. While library automation developers always ask librarians for feedback in what they would like to see in the next iteration of their products, FOLIO gives librarians and library staff a unique opportunity to truly lead this development, through the many paths by which individuals can participate in developing the FOLIO project. 🐙

continued on page 20