SobekCM Digital Repository : A Retrospective

By Mark Sullivan (12/11/2014)

As 2014 draws to a close, the time is ripe for a retrospective on the accomplishments of the SobekCM community over the last twelve months. This year has been busy and the growth of our community continues!

Community Growth

More than a hundred institutions are now hosting materials in SobekCM digital repositories either individually or within collaborative projects. As this software was originally born in central Florida and supports the collaborative Digital Library of the Caribbean, among others, it will come as no surprise that many of the contributing institutions are in the eastern portion of the United States and the Caribbean. But we are also starting to have quite a presence in Europe, Hawaii, and recently added an archive in Pakistan to our installation list.
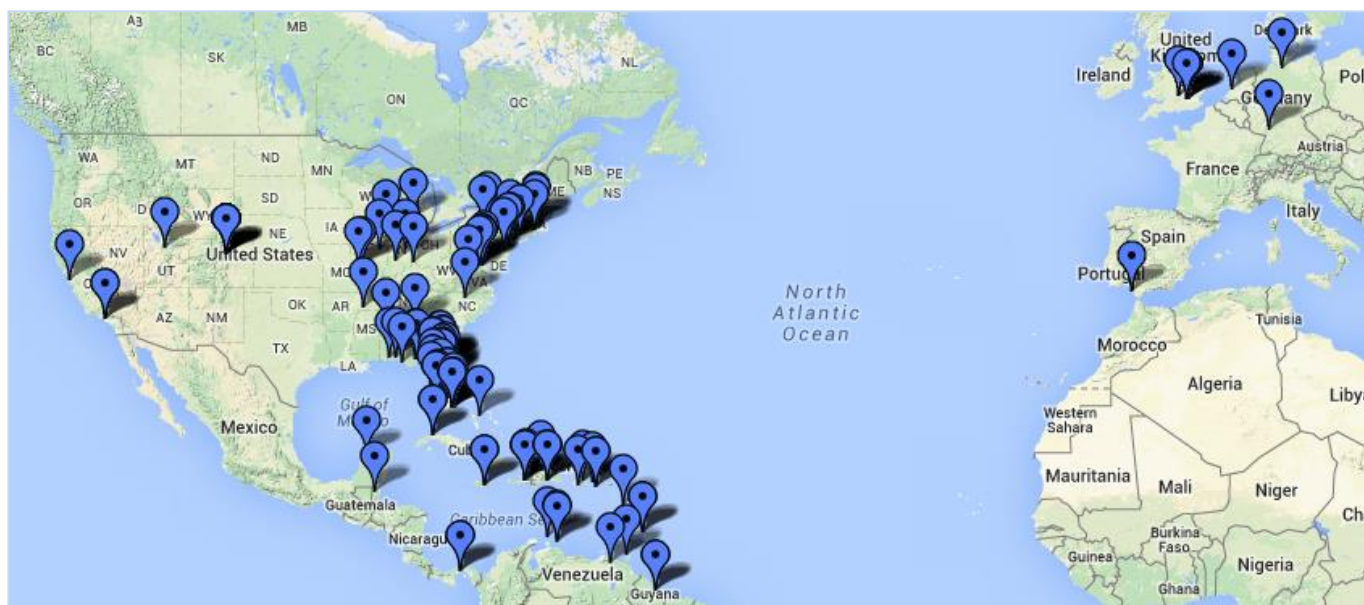

Figure 1: Some of the institutions hosting content in SobekCM instances

The number of known installations of the SobekCM has more than quadrupled in the last eighteen months as more institutions have begun to utilize the open-source software for their individual collections and now includes city archives, museums, and libraries. In addition, several universities have begun to develop either within the main code base, or have written their own digital repository wrappers around the SobekCM core (something we will be encouraging even more with upcoming changes).

All of the custom schemas have already moved to the new SobekRepository.org site and all the help pages should move there over the next couple weeks. We will be discussing building steering committees and core developer groups over the next several months to help encourage community feedback and guidance. This mission and information about those groups will appear on the SobekRepository.org site in the upcoming months.

Major Updates

Over the last twelve months several new versions of the software have been released, which included many new features and some major rewrites. Perhaps the biggest addition for newcomers has been the new installer which was first released early this year. This greatly simplified the process of installing new versions of the software and performs requirements checks and walks users through a much simpler installation and configuration process.

HTML5 Compliant

Nearly every line of HTML code was examined and refactored last year as we made the leap to the future of the web with HTML5. At the same time, all of the stylesheets were examined and far more ways to customize the look and feel of your library were enabled throughout that process. During the major overhaul of the look and feel of the library, the old tabs were retired for sleek new main menu systems throughout the system.

Performance Tuning

The system's already excellent performance was augmented through a number of changes in the way the pages are created on the server and by minifying the amount of code necessary to download to render the pages in your browser. This allows quicker performance especially on mobile devices and in locations with slower internet access. The SobekCM library outperformed other competitors on similar pages both in speed of response and overall size of the pages.

New Features

Many new features were added or significantly updated over the last twelve months by our developers. These main new features can be lumped into categories based on the role of the users that most benefit, although naturally they all benefit your public users and researchers in the end.

Public Users

Much of the focus during the last year focused on incrementally upgrading the overall look and feel of the system and many small features and changes appeared during that process.   Some particularly large changes and additions did occur during this process and are highlighted below.

A new zoomable server was implemented and is packaged with the SobekCM software, allowing for smooth zooms and a cleaner, more responsive feel by utilizing IIPImage server and SeaDragon technology. An excellent full-screen mode is included in this new viewer as well. We continue to fine tune the process and creation of JPEG2000 images to maximize the appearance in this new viewer.

Some new hover functionality was added last year as well. When a user hovers over the subcollection buttons on the collection home pages, more information about the collections appears in a small hover window. This allows users to get a quick description of the collection, before clicking into it.
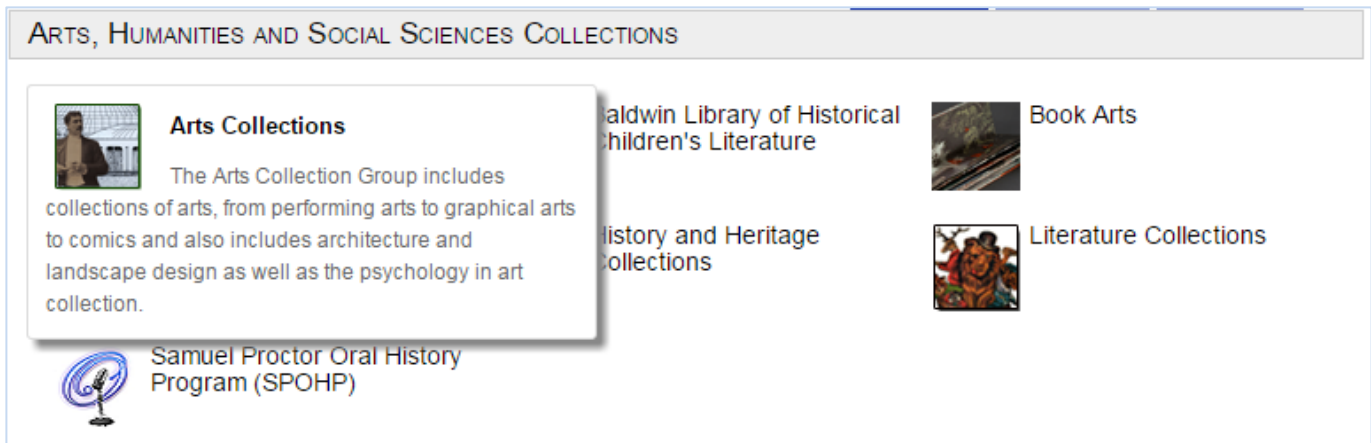
Figure 2: Hovering over the collection buttons makes a small hover with the description appear

Similarly, when a user hovers over an item in the thumbnail view you will see more of the citation for the item in question. This allows users to use the thumbnail view more efficiently and still discover the details about the items within your collections.
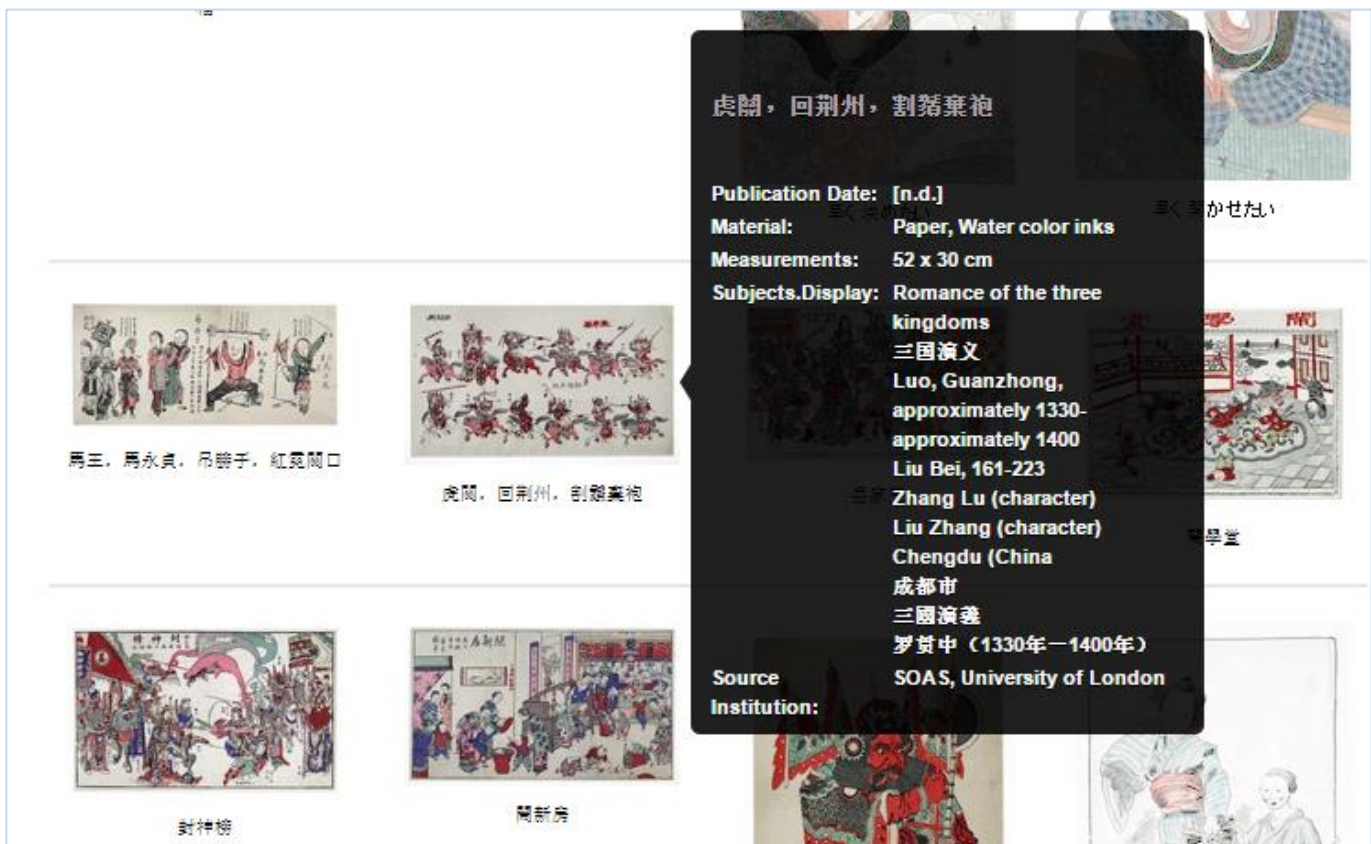

Figure 3: Hovering over a thumbnail after a search or browse in thumbnail view shows more details about the item

Submitters

Some features are particularly useful for users that submit materials to a collection online. Much of this work was a response to requests from institutions and represent completely new features or upgrades to recently added features.

We added the ability to upload multiple files simultaneously to the online item submission interface in the last year. This was a much requested feature; previously each file had to be uploaded individually.

Figure 4: An example of uploading multiple files simultaneously.

The metadata online tools and templates continued to expand. New elements were added to support special electronic theses and dissertations metadata and to support entering and editing DarwinCore metadata for specifying species information. The online Quality Control tool, used to assign structural
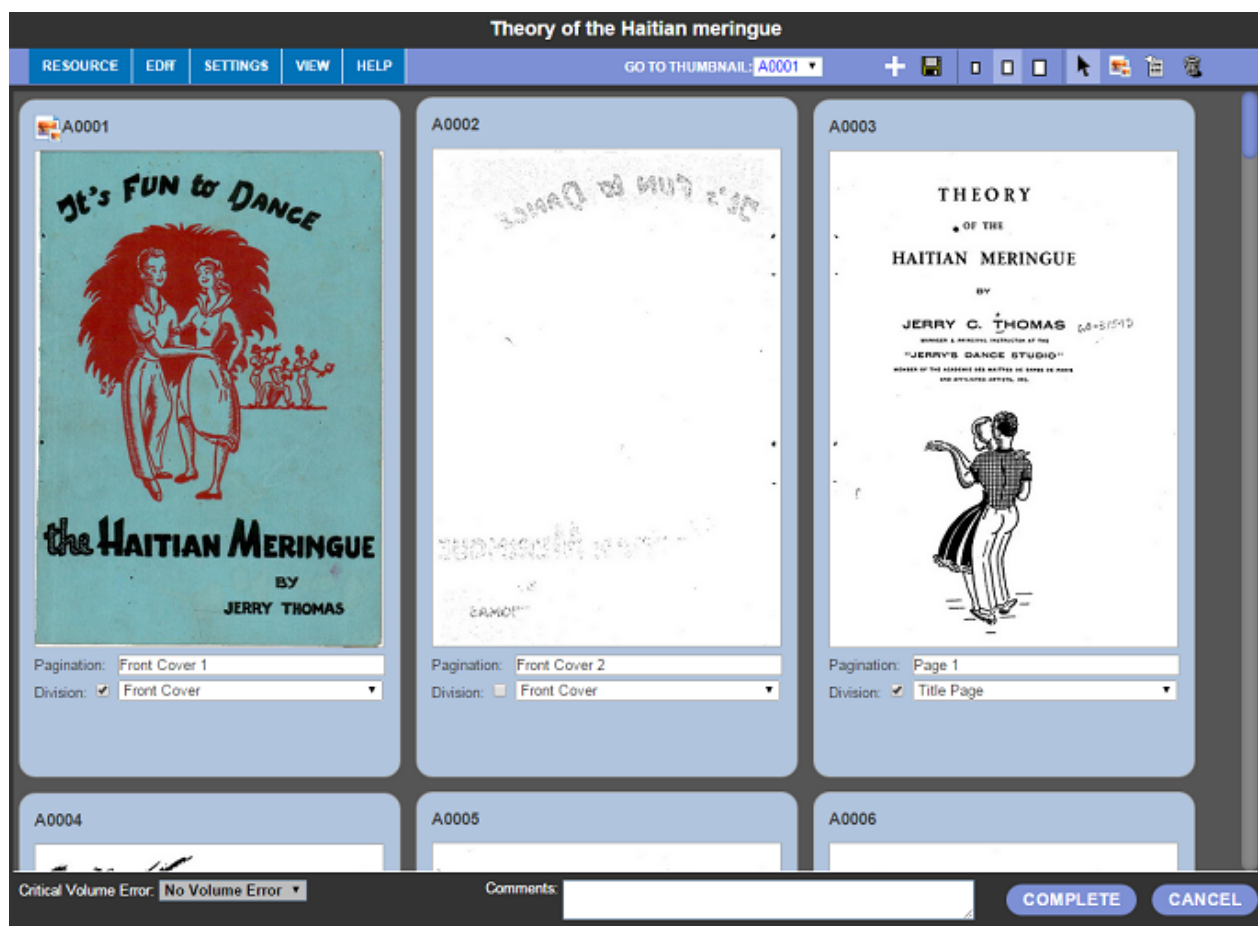
Figure 5: An example of using the online quality control tool to create structural metadata

metadata to the digital resource online, reorder pages, and generally perform image quality control, continued to be refined and began being used as a primarily tool by many submitters to create their table of contents (METS files) online.

Simple manage menus were added and appear under the MANAGE tab when logged on users are viewing an item they have permission to change. Menus were added at both the item and title views, to allow users to clearly see what actions they can take on the items online. This augments the internal header at the item level, while providing clarification on each possible action.
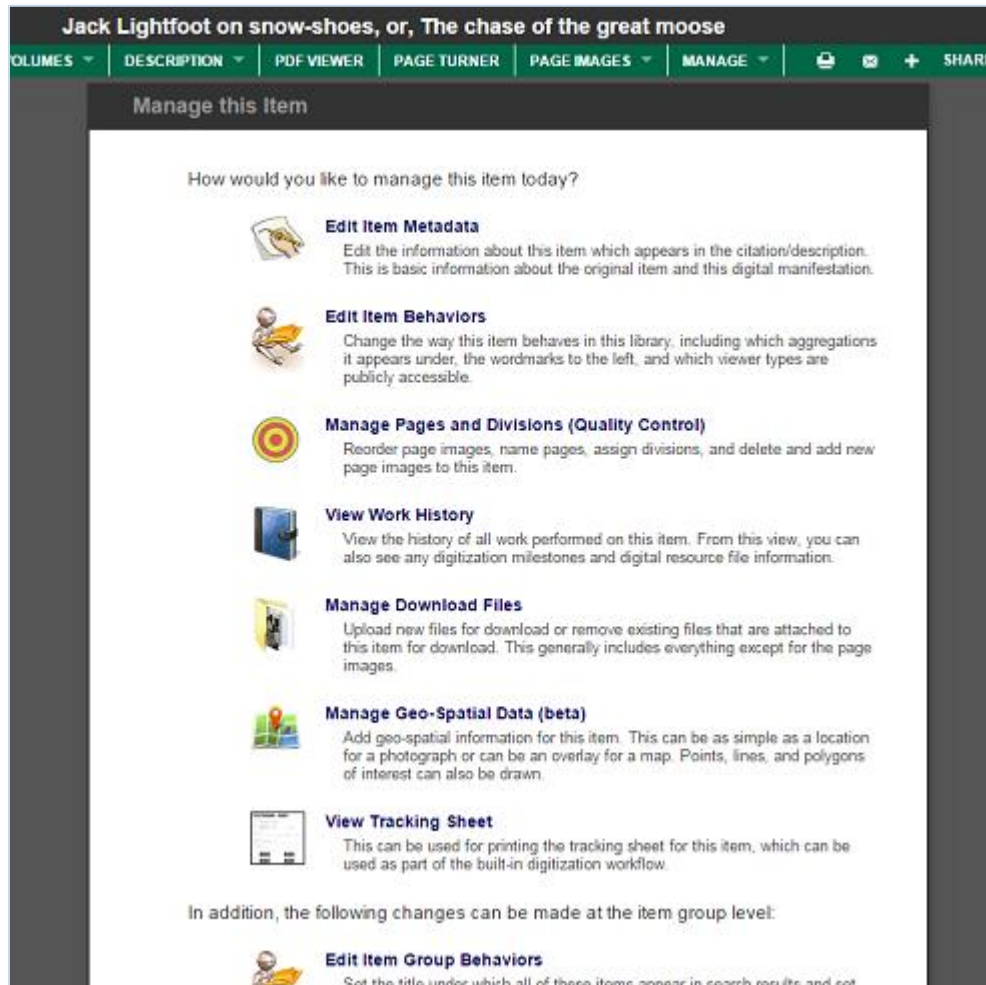
Figure 6: New *Manage this Item* screen for logged on users with access to manage the item

Collection Managers and Curators

Online management of collections and the items in the collections finished a major overhaul over the last year as well. The work on the online curator/admin screens completed early this year. If you have not had a chance to use these new tools on your collections yet, be sure to register to manage them and try them out. You can also see them in use (and many other features) on our SobekCM YouTube page.

Figure 7: One of the many new collection curator/admin screens allowing every facet of the aggregation to be managed

We began a revamp of permissions at the item levels and embargos as well. The system now supports setting IP restrictions over a number of IP ranges and setting an embargo date online through an easy-to-use interface. And the new builder will automatically unembargo the items once the date is past and send an email to the collection manager to let them know which items recently came off embargo. (See figure 8 on the next page).
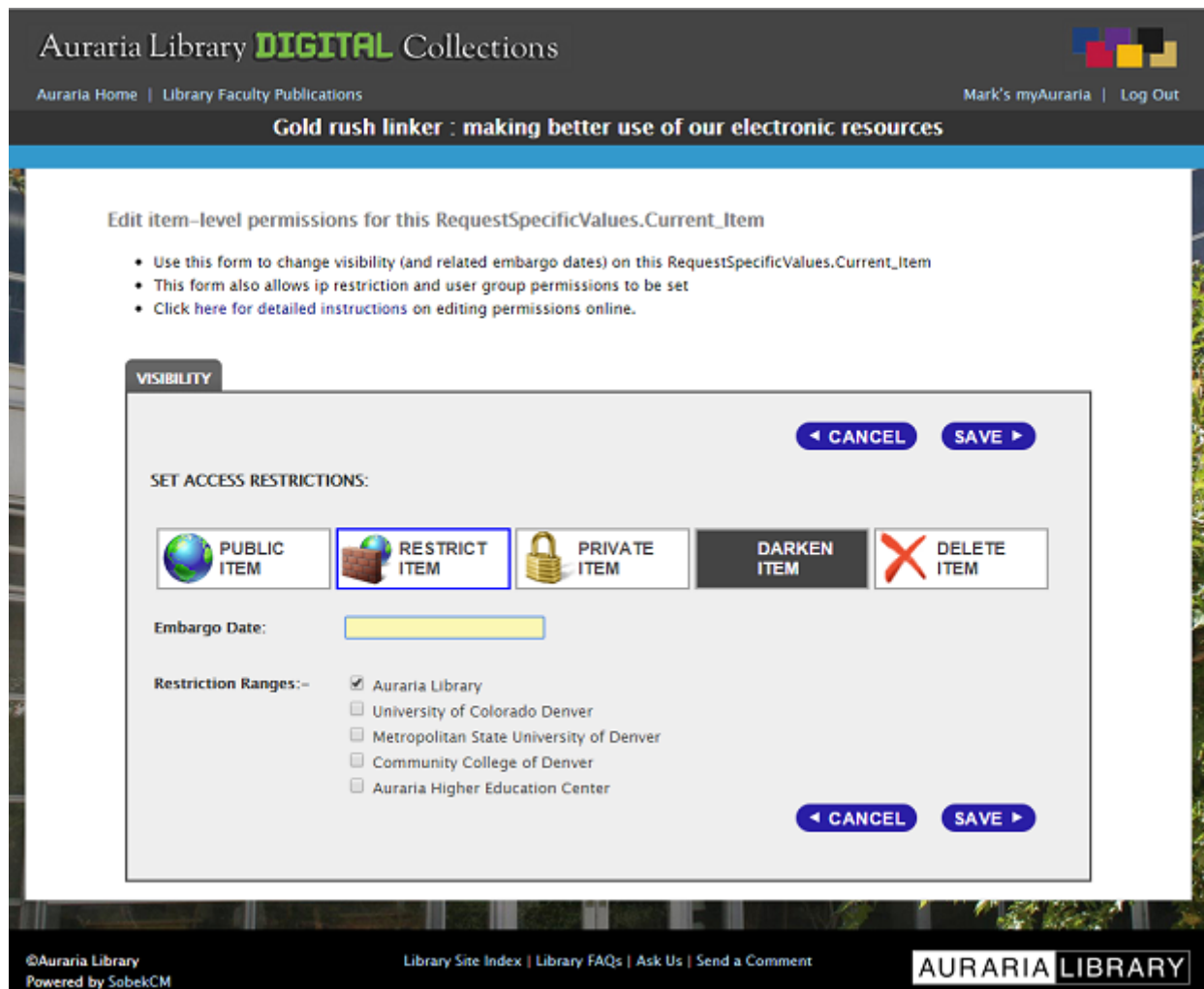
Figure 8: An example of setting embargo and restricting to particularly IP restriction ranges until then

Beta Features

Many features were rolled into the web application as beta features over the last year and are expected to finish development over the next couple months. The map edit feature for adding geographic information to your resource visually utilizing google maps was largely completed and is in final testing. New support for datasets has been rolled out and we are examining the tools necessary to allow users to upload their datasets and describe them online to make the support more accessible. Localization files have been created to allow the system to easily support a number of new languages and allow for further growth in the future. We expect over the next couple months to finish implementing the new localization system. Finally, support for Shibboleth has been normalized within the system, so if you use Shibboleth within your institution or consortia, the system can now support single signon for your users.

Architecture Changes

A number of architectural updates occurred early last year and several more are under way for next year. The builder was updated to support multi-instance environments for those institutions that run a number of different instances and want to use one builder to keep them all current. In addition, changes in code allow for new item-level and aggregation-level viewers to be easily added to the code

base, taking advantage of a number of ways they can define their overall behavior for the system. The builder underwent a radical refactoring to support increased modularization and customization by instance recently as well. You can expect more information about this on an upcoming technical brief.

We have just finished major code reorganization within the web application in preparation for the release of SobekCM 2015 next year, which revolutionizes the internal architecture of the system. The new system will have two clearly defined portions, the UI (user interface) and the SobekCM engine behind it. The SobekCM engine is really a collection of micro-service endpoint definitions, available via REST APIs for the user interface. This division will allow for a much greater ability to customize the system, or create brand new front ends in any programming language. For more information, read the last technical brief on the subject.
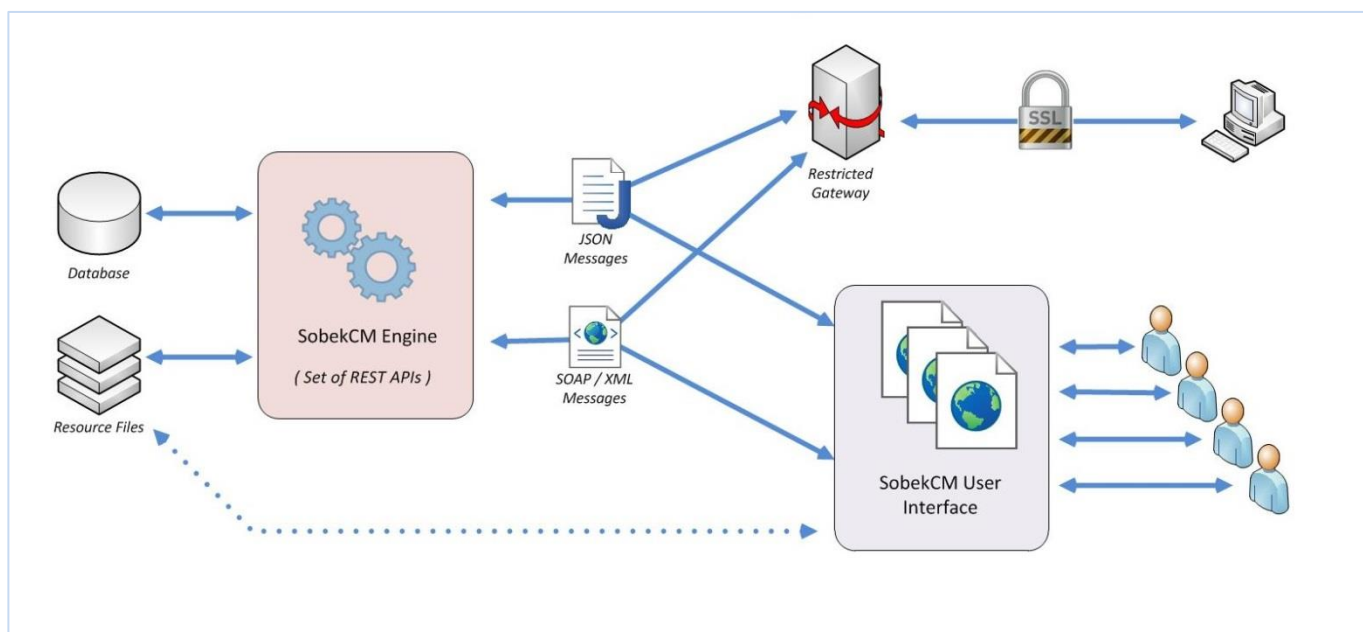

Figure 9: An overview of the separated architecture for SobekCM 2015

For more information, and to keep up with the latest on Sobek Digital, visit our news page at https://sobekdigital.com/news .