

5-6-2019

Building a Better Book Widget: Using Alma Analytics to Automate New Book Discovery

Travis Clamon

East Tennessee State University, clamon@etsu.edu

Follow this and additional works at: <https://dc.etsu.edu/etsu-works>



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

Citation Information

Clamon, Travis. 2019. Building a Better Book Widget: Using Alma Analytics to Automate New Book Discovery. *Code4Lib Journal*. (44). <https://journal.code4lib.org/articles/14371> ISSN: 1940-5758

This Article is brought to you for free and open access by the Faculty Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in ETSU Faculty Works by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Building a Better Book Widget: Using Alma Analytics to Automate New Book Discovery

Copyright Statement

© Travis Clamon. This document was originally published in *code4Lib Journal*.

Creative Commons License

Creative

Commons

This work is licensed under a [Creative Commons Attribution 3.0 License](#).

Attribution

3.0

License

Building a better book widget: Using Alma Analytics to automate new book discovery

Are we doing enough to market newly acquired book titles? Libraries purchase and subscribe to many new book titles each year, both print and electronic. However, we rely on the expectation that users will periodically search our systems to discover newly acquired titles. Static lists and displays have been traditional marketing methods for libraries, but require tedious time and effort to maintain. Without a practical solution for an academic library, East Tennessee State University developed an automated process to generate book widgets utilizing data from Alma Analytics. These widgets are now deployed in our subject guides, website, and on our digital displays. This article outlines the development and implementation of these widgets. We also discuss the challenges we encountered, such as finding image covers and custom subject tagging.

by Travis Clamon

Introduction

Libraries have traditionally promoted new titles utilizing a shelf or display in the library to encourage discovery. Technology in the past decade has enabled the adoption of electronic books as well as growth in online education. While the idea of new title discovery is still important, the traditional method relies on an outdated expectation that users solely rely on the physical presence of the library. This method of delivery and promotion needs to adapt digitally and incorporate electronic and physical books as a whole.

While discovery systems handle virtual searching and browsing of books, they do not enable new title discovery out of the box. Some librarians often resort to creating static lists of curated titles. However, current e-Book ownership models have complicated matters of keeping displays and title lists current. The number of additions and deletions often reaches the thousands as eBook vendors release updates throughout each semester. To have reliable data, we need to be able to retrieve automated data to keep titles accurate.

At East Tennessee State University, we have developed a program that enables automated new title discovery utilizing Alma, our current library management system.

Specifically, we used Alma Analytics, the reporting component of Alma. These reports, which contain new title data, refresh each day when Alma runs indexing. This article will outline the development and implementation of these widgets.

Project Background and Goals

In the past two years, we have made multiple efforts in addressing the goal of new title discovery. The initial idea of automated book sliders occurred when the library was redesigning the website in 2017. One primary task of the redesign was to create uniform subject guides for all of our academic programs. One initiative of the subject guides was to make users aware of the new books in their subject or program. Our efforts of creating a book slider for each subject guide relied on separate Alma analytics reports and scripts for automation. Shortly after the website went live, the library also implemented a new print book widget on our digital display. We soon realized the workflow of each widget was becoming redundant and could not scale to include multiple subjects, formats or templates. Implementing new data columns or changes required editing numerous analytic reports and scripts. As the number of individual book widgets and interest from librarians increased, we decided in late 2018 to rewrite our code to introduce new features and eliminate redundancy.

The development of the project happened as time allowed by the electronic resources librarian. A total time is unavailable, but we estimate that 50 hours went into developing the new version. Before the rewrite began, we identified technical and user requirements that would satisfy existing workflow issues:

- Implement widgets on subject guides, webpages, and digital displays.
- Multiple templates to choose from such as slider widgets and generated table lists. New templates are added without additional configuration.
- The ability to pull data on new title information (print and electronic) from a minimal number of Alma Analytic reports.
- Create a widget based on a particular subject, format, or physical location.
- Ease of use; Limited technical expertise needed to create a widget

High-Level Overview

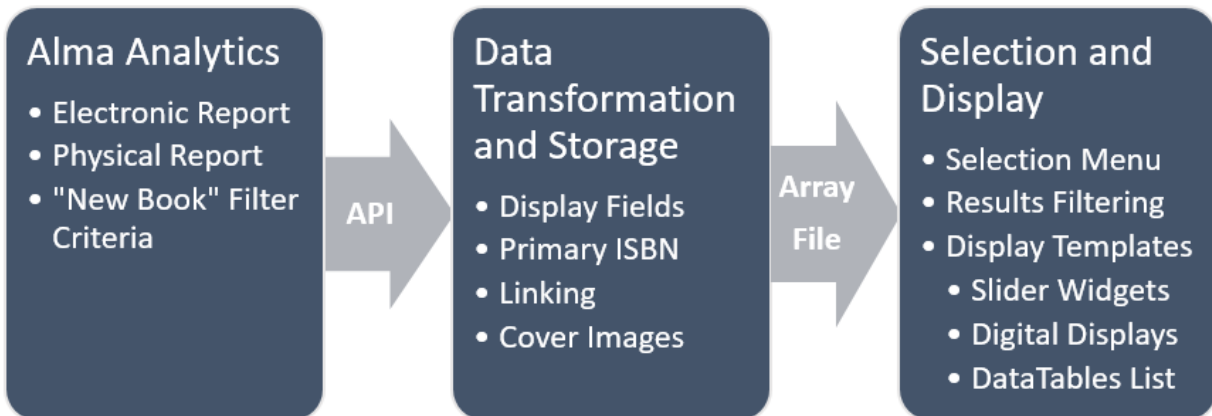


Figure 1. High-Level Overview

We divided development into three distinct processes:

1. Alma Analytics: Reports have been created that warehouse the new book data. We then retrieve the data using the Alma Analytics API.
2. Data Transformation and Storage: We transform various data fields and store them into an array file.
3. Selection and Display: Accepts the array in step 2 along with additional parameters to create the widget. A selection menu allows librarians to preview their widget and receive an embed code for display.

Section 1: Alma Analytics

The first step in the process is creating an Alma Analytics report. When creating a new report, a user will first select a subject area. Since our goal is to include both print and electronic books, we created an analytics report for each format in order to capture all necessary data fields. The print report comes from the “Physical Items” subject area, while the electronic report comes from the “E-Inventory” subject area.

Report Fields

An important next step of the process is identifying the data fields needed for the project. This is dependent upon the intended uses of the widgets and the information you want to provide. Based on our needs and previous experiences, we were able to identify the following fields in each format:

Data Field	Physical Items Report	E-Inventory Report
Title	Bibliographic Details -> Title	
Author	Bibliographic Details -> Author	
ISBN	Bibliographic Details -> ISBN (Normalized)	
Publication Date	Bibliographic Details -> Publication Date	
Upload Date	Physical Item Details -> Creation Date	Portfolio Creation Date -> Portfolio Creation Date
Subjects	Bibliographic Details -> Subjects	
LC Classification Group	LC Classifications -> Group1 LC Classifications -> Group2 LC Classifications -> Group3 LC Classifications -> Group4	
Call Number	Holding Details -> Permanent Call Number	N/A
Location Code	Location -> Location Code	N/A
Location Name	Location -> Location Name	N/A
Collection	N/A	Electronic Collection -> Electronic Collection Public Name
MMS_ID	Bibliographic Details -> MMS Id	

Table 1: Alma Analytics report filters for physical and electronic Items

Report Filters

Report filters will define the overall target collection of books. At this step, we had to define what our library considered “new” for electronic and print formats. Book purchasing and subscription changes will vary in each library. We also imposed a publication date range to ensure that we do not feature older materials. We tested various year ranges to get an idea of the amount of results we wanted to store in a dataset. East Tennessee State University defined the following criteria for each report:

- Publication Date must be between the past three years up until the future year.
- ISBN is not null. The ISBN is used for linking and cover image purposes.
- Portfolio or Item creation date must be during the past 180 days.
- Item must be available or activated for online use.

- Specification of Item locations and/or electronic collections are optional. We excluded Government Document books using this filter since they were not relevant.
- If the location contains more than books, a material type filter is required.

Retrieving Report Data from Alma Analytics API

Once the report was created and we felt satisfied with the selection criteria, we used the Ex Libris Alma Retrieve Analytics report API

(<https://developers.exlibrisgroup.com/alma/apis/docs/analytics/>) to receive and store the data from both analytics reports. The API requires a key along with the path to the analytics report. If more than 1000 results are returned from an individual report, a resumption token allows you to receive the remaining results.

One issue we encountered when receiving the data is that the column names are generic (column1, column2, etc.). The column numbers returned did not match the column order in the Alma Analytics report display either. Also, we noticed fields such as the LC classification groups had different column numbers between the print and electronic responses. Fortunately, these columns remain the same after repeat API calls. We resolved this matter by creating an array for each report in our script that cross-matches the response column number to a more appropriate variable name.

Section 2: Data Transformation and Storage

Once we have the data from Alma Analytics, some additional transformation is required to get the data into a usable format. From the API response, we convert the data to an XML Element and process each column using a foreach statement. Most of the fields were a direct copy over into the final array, but others needed some additional work:

- **Title:** We noticed some of our book titles had all upper case letters, mostly due to one particular vendor's MARC records. In order to make a consistent appearance, we used the PHP `mb_convert_case` function and specified a `MB_CASE_TITLE` conversion. Slashes were also common at the end of each title. We used regex to remove the slashes.
- **Publication Date:** This field varied by record and sometimes included copyright symbols, periods, and brackets. We used regex again to remove the invalid characters and leave only digits remaining. This was required so we could sort the publication date correctly.

ISBN

The ISBN is used for cover image and linking purposes. Since most records contain multiple ISBN numbers, we needed a way to determine which ISBN would return a cover image from our discovery system, Ex Libris Primo. We first started using the first ISBN provided, but quickly realized that it did not always return a valid image.

We resolved the issue by breaking up the ISBN's into an array. Each ISBN assembles as an URL and then a cURL call checks the file size of the URL. The Primo service returns a small 1x1 white pixel if a cover image is not available. We determined the white pixel returned by Primo was consistently 86 bytes. Once the first ISBN came across that had a file size other than 86 bytes, the loop would end and the ISBN would be become the preferred value.

At this point, we analyzed the data again and determined that around 1% of print books and 5% of electronic books were unable to retrieve a cover image from Primo. We queried these ISBN's with Google Books and were able to retrieve a few more cover images. This resulted in us modifying the code to do an additional call using the Google Books API (<https://developers.google.com/books/>) only in cases when a cover image is unavailable from Primo.

We quickly noticed that adding the full array of ISBN's along with the Google Books API made the import process much longer and would sometimes timeout. In order to reduce the amount of repeat cURL calls every time the import process occurs, we started storing the cover image URL, current date, and the corresponding MMS_ID (Alma record ID) in a separate array file. We scan this array before initiating a cURL call. The array retains data between imports. As time continues and the array grows, we will clean out older records by filtering the date stamp.

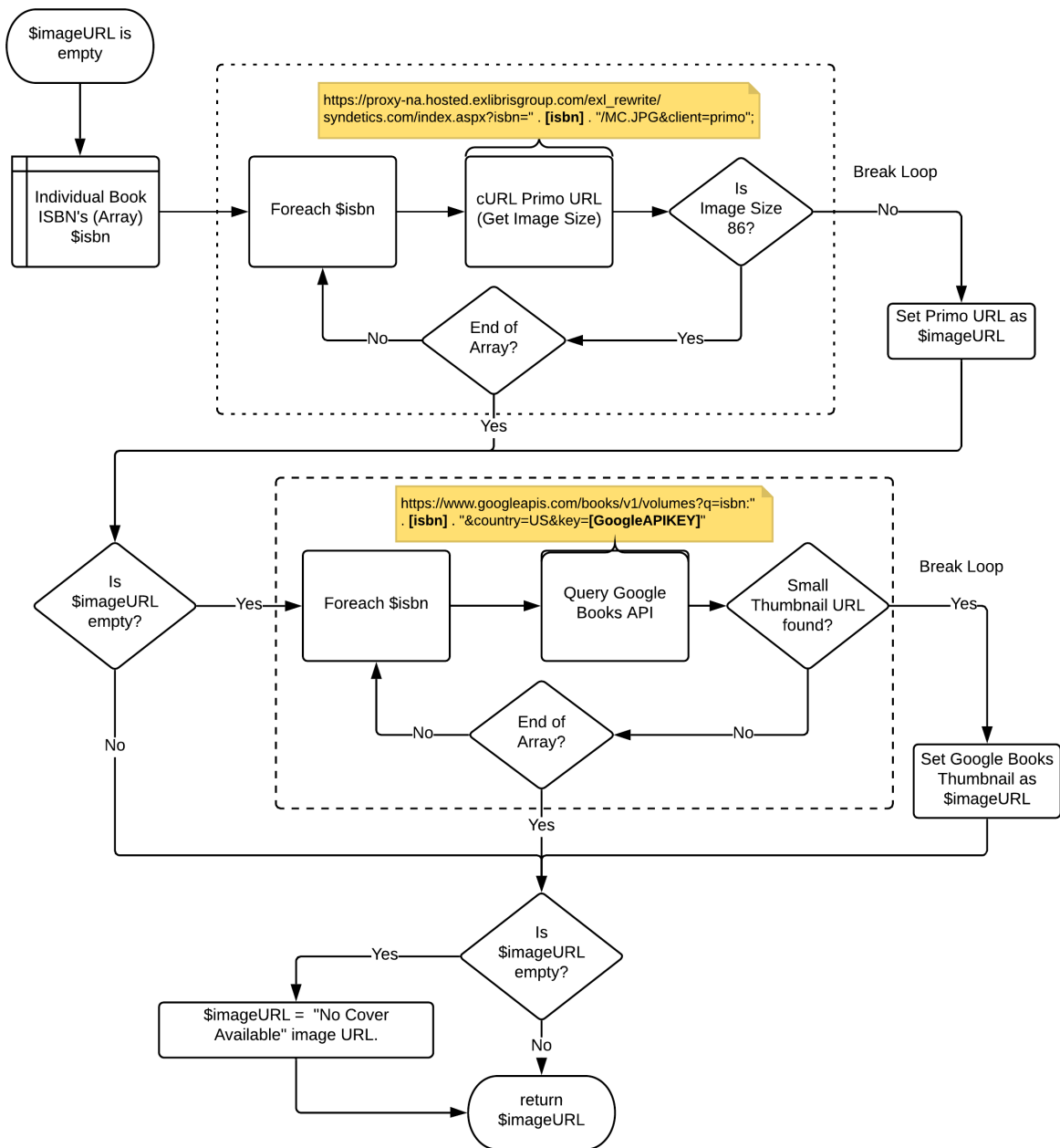


Figure 2. Cover image retrieval process using Primo and Google Books API (see [code](#))

Custom subject tagging

Another relevant section was finding a way to assign custom subject tags to books. These tags correlate to the subject guides available on the library website. Due to the variations of vendor-supplied MARC records, we determined keyword matching requires the LC group classification, title, and subjects columns. A case statement contains all

fifty-seven custom subject terms for the university. We made matches for each subject tag to the corresponding Library of Congress classification terms. If a match does not occur by LC Group, we attempt to do a keyword match using the subjects and title columns from the report. In a similar fashion, custom keywords are also matched using a case statement. These case statement files are stored separately as include files in the subjects folder and are available to reference from the GitLab repository (<https://gitlab.com/clamon/alma-book-widget>).

Links

Subject guide widgets have links attached to book covers, along with titles in a list view. Instead of linking directly to an eBook or a detail page for a physical book, we decided to send all requests to our “Get It @ ETSU” link resolver. We decided on this approach for statistical and maintenance reasons. The primary ISBN is the parameter the link resolver uses to locate the title. The other parameter we used is the service id “sid” parameter. We assigned a default value of books_widget. This value allows us to retrieve usage statistics using the link resolver object in Alma Analytics. The link resolver can also provide the most current availability information on a book, rather than the 24-hour index window in Alma Analytics.

UNIVERSITY LIBRARIES
EAST TENNESSEE STATE UNIVERSITY

Google, making information accessible : the search engine that changed the world /
Fastré, Guillaume, author.
ISBN: 2808005792 , 2808005792 , 9782808005791

Online access >

TOP

SEND TO

VIEW IT

LINKS

GET THIS

TAGS

Send to

View It

Sign-in to Get This Sign in

Full text available at: [EBSCO eBooks Business Collection](#)

Additional services

[Report Problems](#)

Report a problem with this resource to Sherrod Library Technology & Content Services. Average response time: 1 business day

Feedback

https://etsu-edu-primo.hosted.exlibrisgroup.com/primo-explore/openurl?isbn=2808005792&sid=books_widget

Primo Services Page URL Book ISBN Source ID

Figure 3. Sample URL linking to the Primo services page using ISBN and Source ID parameters.

Request Date	Title	Author	ISBN	Source
3/7/2019	Principles and practice of pediatric infectious diseases /		9780323401814	books_widget
3/11/2019	Meyler's side effects of drugs : the international encyclopedia of adverse drug reactions and interactions /		9780444635938	books_widget
3/12/2019	3D integration in VLSI circuits : implementation technologies and applications /		9781138710399(Hardback)	books_widget
3/13/2019	Applied hierarchical modeling in ecology : analysis of distribution, abundance and species richness in R and BUGS /	Kéry, Marc, author.	0128013788;9780128013786	books_widget
	Images of the art museum : connecting gaze and discourse in the history of museology /		3110341360	books_widget
	Mastering the Nikon D850 /	Young, Darrell, 1958- author.	1681983729	books_widget
	Numerical simulation in hydraulic fracturing : multiphysics theory and applications /	Shen, Xinpu, 1963, author.	9781138029620(Hardback)	books_widget
3/14/2019	Basic & Clinical Pharmacology, 14e		1259641155	books_widget
	Clinical pharmacology /	Brown, M. J. author. (Morris Jonathan),	9780702073281	books_widget
	Goodman and Gilman's Manual of Pharmacology and Therapeutics	Hilal-Dandan, Randa., ath.	9780071792882.	books_widget
3/15/2019	Braunwald's heart disease : a textbook of cardiovascular medicine /		9780323463423	books_widget

Figure 4. Sample Alma Analytics link resolver report of widget click usage.

Data storage and automation

Once all the data processes, the final output consists of a multidimensional array stored as a file inside the export folder. The array size varies, but it typically contains between 1,000 to 5,000 book titles. We chose the array file over a database due to low maintenance and its compatibility with the DataTables plugin. This process runs daily using a cron job on our server. To confirm that the call and import were successful, we inserted a Slack webhook (<https://api.slack.com/incoming-webhooks>) to notify us of the status each day.

Section 3: Selection and Display

We developed the selection and display component separately from the transformation part. This component accepts the multidimensional array file created after transformation. When generating a widget, three parameters are required: filter type, filter query, and template. These parameters are imported from PHP `$_GET` variables in the URL. The filter type correlates to an individual data column inside the multidimensional array. With our current setup, we can filter data on types such as title, author, publication year, location code, subject, or type (physical/electronic). The filter query is the search keyword. We created a function that accepts these values, searches the array, and calls upon the required template to generate the widget HTML.

```
1  function createSlider($array_source, $filter, $query, $template) {
2      $results_array = array();
3
4      //find results
5      foreach ($array_source as $key => $val){
6
7          if(is_array($val[$filter]) === true) {
8              foreach($val[$filter] as $term) {
9                  if(stripos($term, $query)!= false){
10                     $results_array[] = $val;
11                 }
12             }
13         }
14         elseif(stripos($val[$filter], $query)!= false){
15             $results_array[] = $val;
16         }
17     }
18
19     //Pass results_array to the designated template
20     require "templates/" . $template . ".php";
21 }
```

For ease of use, we developed an interface that assists with creating a widget. We scan the original array to show all possible combinations of data for filter types as well as templates. Both title and author allow free-text input, while the other fields present as options (radio buttons). Once all selections are made, a widget preview frame appears, and embed code is generated.

Example widget selections:

- New books located on the second floor (Location Code).
- New books published in 2019.
- New books subject-tagged in Nursing, Computing, etc.
- New books in Research Methods (free-text keyword).

Which filter would you like to apply?

title author pub_date upload_date mod_date location_code subjects type

Choose a value for your filter...

online esbb4 esbb3 esbsc esbnw esbb2

Choose a template...

BrightSign (Digital Display) Subject Guides

Embed Code

https://libs.etsu.edu/books-display/dev/display.php?filter=location_code&que



Figure 5. Internal interface that assists with book selection and widget implementation

Display templates

A template folder retains multiple different widget display options. We had three formats we wanted to begin with: a small sliding widget for subject guides, a full-screen slider widget for our digital display, and a list view for informational needs. From a technical perspective, display.php processes the widget query, and a new JSON array is created with the results. At this point, we use the template parameter specified from the URL to

include the corresponding PHP file (“template”) that is being used. Each template file will then process the results array as needed to create the intended widget. Since most of the development time focused on building the data array, we wanted to utilize open source components for our templates. These plugins are stored in a separate folder for easy updating.

For the sliding widgets, we found a carousel package called Slick developed by Ken Wheeler (<http://kenwheeler.github.io/slick/>). This package allows for single and multi-item sliders that are responsive to mobile devices. For the subject guides, we used a multi-item carousel and specified three different responsive configurations. The setup allowed us to show seven book covers at a time for desktop displays while offering a mobile and tablet friendly display of two to five sliding book covers. HTML title tags were added containing book titles for accessibility purposes. The slider automatically rotates every few seconds and the arrow buttons allow for manual control.

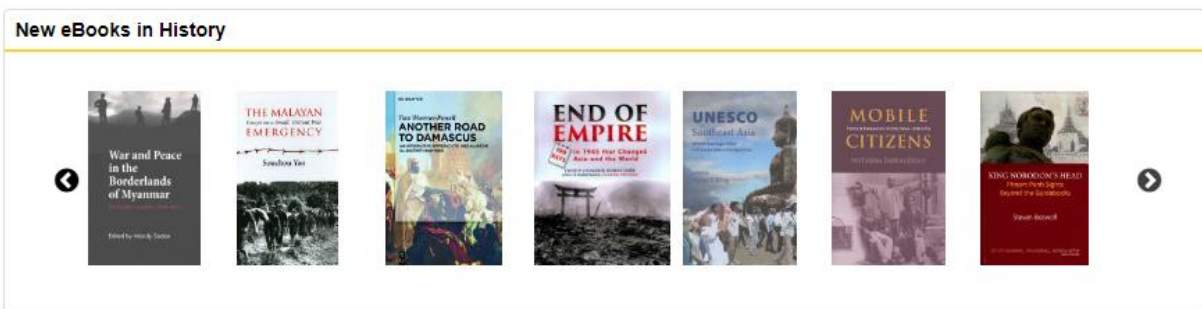
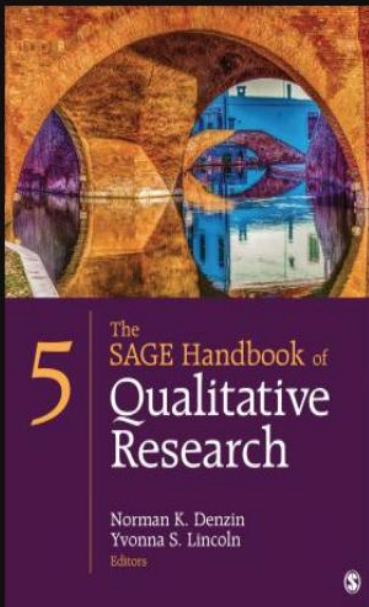


Figure 6. Slider widget embedded on the history subject guide (<https://libraries.etsu.edu/research/guides/history/books>)

We encountered some initial difficulty importing this widget onto our subject guides. We tried first using the remote script option, and while the slider would initially appear, any screen resizing would cause the widget to disappear. We determined there was a jQuery conflict between LibGuides CMS and the Slick package. Instead of using the remote script function to pull in the widget, we ultimately decided to use an iFrame. Since HTML code is necessary, instructions and code are provided at the selection menu along with a copy to clipboard function.

For the digital display, we utilized the single item carousel and customized the CSS to fit the display appropriately. In regards to visual design, we chose a dark background with white text that conveys title, author, and location. Since the display constantly rotates throughout the day, we configured BrightSign (our digital display hardware) to refresh the webpage every 15 minutes. When it refreshes, the script will randomly choose a new order of titles to give a new and fresh appearance.

New Books at Sherrod



The Sage Handbook Of Qualitative Research

Norman K. Denzin, Yvonna S. Lincoln

ETSU Sherrod Books (2nd Floor) (H62 .H2455
2018)

Figure 7. Slider widget embedded on our digital display inside the main entrance.

The list view is useful when providing an overall list of new books. We were able to implement the DataTables plugin (<https://datatables.net/>), which enhances HTML tables with little effort. DataTables accepts our converted JSON array and displays the appropriate fields using a custom display configuration. The subject and type (electronic/physical) were configured as dropdown sorting menus. We had to do some additional work to get the subject dropdown to work and sort alphabetically by breaking up the array of tagged subjects. To reduce loading time, we set `deferRender` equal to `true` so data is only loaded for the current page of results and not the entire book array. This is important when you are displaying book covers so the browser does not request thousands of images at once.

We also implemented button plugins to allow for Excel data exports. This feature is useful for end users who want an offline copy, but as well for librarians who want to send a copy to interested faculty. We also explored PDF exports using the PDFmake plugin, but we encountered issues with book url's splitting into multiple lines on the PDF. Adobe Reader would only recognize the first line of the URL when clicked, and therefore would result in a bad request. This list is available on our main books page and linked from our display widgets on our subject guides.

Browse New Books

Search:

Subject:

[Export Titles to Excel](#)

Cover	Title	Publication Date	Collection
	A Crisis In Swiss Pluralism : The Romansh And Their Relations With The German- And Italian-Swiss In The Perspective Of A Millennium Author: Billigmeier, Robert Henry.	2016	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Dictionary Of Family History : The Genealogists' Abc Author: Scott, Jonathan.	2017	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Field Guide To Spiders Of Australia Author: Whyte, Robert, 1955- author.	2017	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Focus On Sustainable Supply Chains And Green Logistics. Author: Emerald Group Publishing Limited.	2016	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Fresh Look At Potential Output In Central, Eastern, And Southeastern European Countries Author: Podpiera, Jiri, 1976-	2017	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Grand Adventure : The Lives Of Helge And Anne Stine Ingstad And Their Discovery Of A Viking Settlement In North America Author: Ingstad, Benedicte, author.	2017	Online Resource - EBSCOhost Academic eBook Collection (North America)
	A Higher Loyalty : Truth, Lies, And Leadership Author: Comey, James B., Jr., 1960- author.	2018	ETSU Sherrod Books (3rd Floor) (HV8144.F43 C66 2018)
	A History Of Islam In Indonesia : Unity In Diversity Author: Kersten, Carol, author.	2017	Online Resource - EBSCO eBooks Religion Collection
	A History Of Mexican Literature	2016	ETSU Sherrod Books (3rd Floor) (PQ7111 .H58 2016)
	A Most Elegant Equation : Euler's Formula And The Beauty Of Mathematics Author: Stipp, David, author.	2017	ETSU Sherrod Books (4th Floor) (QA255 .S75 2017)

Showing 11 to 20 of 1,340 entries

Previous 1 **2** 3 4 5 ... 134 Next

Figure 8. List widget utilizing DataTables to display the entire array of new books (<https://libraries.etsu.edu/research/books/new>)

Limitations

As designed, a widget only supports one filter parameter. There may be a future requirement to support multiple filters, primarily if a large repository of books exists in the Alma Analytics report. We plan to wait and see how the widgets are used before deciding to rewrite the code.

Subject tagging is another challenging part of the project. It is important that appropriate subjects matching occur. This ensures our users are seeing all new relevant titles in their subject area. We plan to keep expanding the keyword list and look into the Alma

configuration to see why specific titles are not inheriting LC Group classifications in Alma Analytics.

For electronic books, we determine whether an item is new or not by the portfolio creation date. In some cases, when we have to reload an entire electronic collection in Alma, this will cause all of those titles to appear as new again. So far, this has not been an issue, but we currently do not have a workaround devised.

During the cover image retrieval process, it should be noted that the Google Books API has a daily quota limit. We were assigned a limit of 1,000 calls per day. We have made every effort to only call the Google API when needed, but if a large load of new books comes through, it may exceed the limit. You can request an higher quota, but will have to provide more information about your project to Google.

Conclusion

The development of this widget tool has been a worthwhile project for the library. The digital display widget on our main lobby has captured a lot of attention by both patrons and librarians. Our feedback for the display has been positive overall, and we have made minor adjustments such as extending the timer between title changes to allow for longer recognition of the item location and call number (if physical).

We updated subject guides to include the new widgets and were able to eliminate over 50 old Alma Analytic reports. Performing results filtering in the script rather than in individual Alma Analytics reports allows for easier modifications and keyword updates. To capture eye attention to the sliding widgets, we positioned them at the top of the template and enabled auto scrolling. An example of our subject guide widget is available at <https://libraries.etsu.edu/research/guides/history/books>.

While the intended purpose of the project is to highlight new books, there is potential for other uses such as highlighting a special collection. All that would be required is modifying the filter criteria in Alma Analytics. As funds allow, we are hopeful that more digital displays are installed throughout the library, and we can focus on marketing books by collection or floor. A goal within the next year is to interview students and faculty and receive additional feedback formally. There may also be potential to expand these displays outside the library, such as in our student center and academic buildings.

The code for this project can be found on the author's GitLab page (<https://gitlab.com/clamon/alma-book-widget>).

Acknowledgements

The author would like to thank co-workers Mark-Hero Agbenu and Destin Hebner for their invaluable assistance and support with DataTables.

About the Author

Travis Clamon (clamon@etsu.edu) is the Electronic Resources Librarian at East Tennessee State University in Johnson City, TN. He holds an M.S. in Information Science from the University of Tennessee and a B.S. in Computer Science at East Tennessee State University. His professional interests include discovery, user experience, electronic resource management, and data analytics.