

Library Analytics Toolkit Report

Project Summary

The Library Analytics Toolkit was originally proposed in the first round of Library Lab funding. The project sought to create an analytics dashboard that would pull library data together in a way that allowed both librarians and patrons to identify and respond to trends and changes in the library's data. The first stage of the project involved researching the types of analytics that libraries across Harvard University and beyond would want to see in such an application. At that time, other analytics projects at other schools were also evaluated to provide inspiration for this project and as a way of gauging what other libraries had found to be the key pieces of information for such a dashboard. This stage of the project culminated in a report (http://osc.hul.harvard.edu/sites/default/files/Analytics_Recommendations.pdf), which made recommendations about the types of data that should be included in the initial version of the Library Analytics Toolkit and ways that it could be expanded in the future.

When the Library Lab agreed to fund the second stage of the project, we moved forward with efforts to build an initial working version of the Toolkit with a focus on visualizing collection and circulation data in a way that would be helpful for collection development professionals and interesting for patrons. This stage culminated in the completion of a working data visualization tool that allows users to explore data about Harvard University Library's collection. The Library Analytics Toolkit is available online at <http://librarylab.law.harvard.edu/toolkit/>. We now look forward to starting working on the recently funded third stage of the project, which will involve extending the tool to encompass more of the Library's collection, including e-resources.

Accomplishments

To date, this project has managed to achieve many of its initial goals and is on track to complete even more of these goals over the course of the third stage of development. Through our initial research phase, we accomplished our goal of determining what librarians are looking for in an analytics dashboard both at Harvard and at other libraries. We also learned from a number of existing dashboard projects that other universities have developed.

In the second phase of the project, we succeeded in creating a working visualization of collection development and use data in a form that makes this information accessible to librarians and patrons alike.

Challenges

Many of the challenges that we encountered during the course of this project related to the fact that libraries use widely divergent tools and techniques to collect data. This is true even of all of the libraries at Harvard University and makes it difficult to create a tool that can ingest and visualize the various formats of data. This was influential in our decision to focus on collection and circulation data in this initial phase of development and will make it more difficult to create a dashboard that visualizes other types of library data.

Next Steps

We have already begun planning the third stage of this project. In this phase, the Toolkit will be expanded to include data about e-resource usage. We also plan to add to the number of items that we are able to visualize in the toolkit by obtaining access to more detailed Library of Congress Subject Heading information. We anticipate that this will allow us to add an additional 200,000 to 500,000 items to the Toolkit.

Beyond this next stage, we hope to be able to add other types of library data to the Toolkit in the future and to be able to offer information about expenditures so that the Toolkit's visualizations can be used to inform purchasing and budgeting decisions. We also plan to post the code for the project so that other libraries outside of Harvard University can develop their own Toolkits.

Confirmation of Code Deposit

We have created a repository on github for the Library Analytics Toolkit: <https://github.com/harvard-lil/Analytics-Toolkit>. Given that the tool is not fully functional we have not yet deposited the code.

Budget Spending

Spending so far has covered initial research in the Harvard Library community to determine needs and the production of a report documenting those needs. Data visualization developers, working with developers on site, then created the prototype. Expenditures were distributed as follows:

Carli Spina	Research and Report	\$1,680
Rosten Woo	Data Visualization	\$8,500
Sha Hwang	Data Visualization	\$8,500
Total		\$18,680

Publicity and Presentations

In connection with this project, we have done several presentations in an effort to gain feedback from other libraries outside of Harvard and to encourage other libraries to consider using our Toolkit when the code is released. We have presented posters on the project at both the 2011 AALL Annual Meeting & Conference and the 2012 ALA Annual Conference (<http://www.slideshare.net/CarliSpina/library-analytics-toolkit-poster>). We also recently offered a 1-hour presentation on the project at the 2012 LITA National Forum (<http://www.slideshare.net/CarliSpina/library-analytics-toolkit-14608730>). At the 2012 ALA Annual Conference, several journals expressed an interest in publishing an article on the project for publication, and we hope to complete such an article for publication in a library technology journal in the coming months.