

Data-Intensive Tools for Modeling and Visualizing Mass Reading

The 'Reading Chicago Reading' Project

Presenters



Robin Burke
Univ. of Colorado,
Boulder



Ana Lucic
Univ. of Illinois



John Shanahan
DePaul Univ.

Additional Personnel and Support

Jennifer Lizak and Michelle Frisque, Chicago Public Library

Lyrasis 2018 Catalyst Fund

NEH Office of Digital Humanities

HathiTrust Advanced Computing Support Program

DePaul University

Ke Feng, Mihaela Stoica, and Nandhini Gulasingam

What is “Reading Chicago Reading”?

- A multi-institution, multi-disciplinary study of Chicago Public Library’s “One Book One Chicago” (OBOC) program
- Uses the OBOC program as a tool for producing new insight into library data
- Works with CPL OBOC circulation data since 2011 and combines it with other sets of data (Census data, CPL events, text measures, social media)
- Recently created a predictive circulation model (see *DHQ*, forthcoming)
- Recently completed a prototype dashboard with Lyrasis support

Predictive Model

Seeking to understand the influence of different factors on branch-level book circulation

- Book content
- Promotional efforts
- Prior popularity
- Neighborhood demographics

Findings

#1: It is challenging to do this kind of modeling

#2: All of the factors are important to different extents

1. Prior popularity
2. Neighborhood demographics
3. Book content
4. Promotional efforts

Some findings about branch-level circulation

- “Live” programming seems to have a small effect on OBOC circulation. Implications for librarians: how and where to use scarce resources (social media promotion vs. book discussion vs. family-friendly “tie-in” events)
- Segregation in Chicago is reflected in its OBOC data. This is esp. important because sponsored mass reading programs aim for new forms of “imagined community.” What books are more likely to spread “evenly” across the city map? (see dashboard) We have found for instance that north branches are more likely to take up a Chicago-set texts, but **south branches do respond differently to book choice** (if not to the amount of CPL programming).

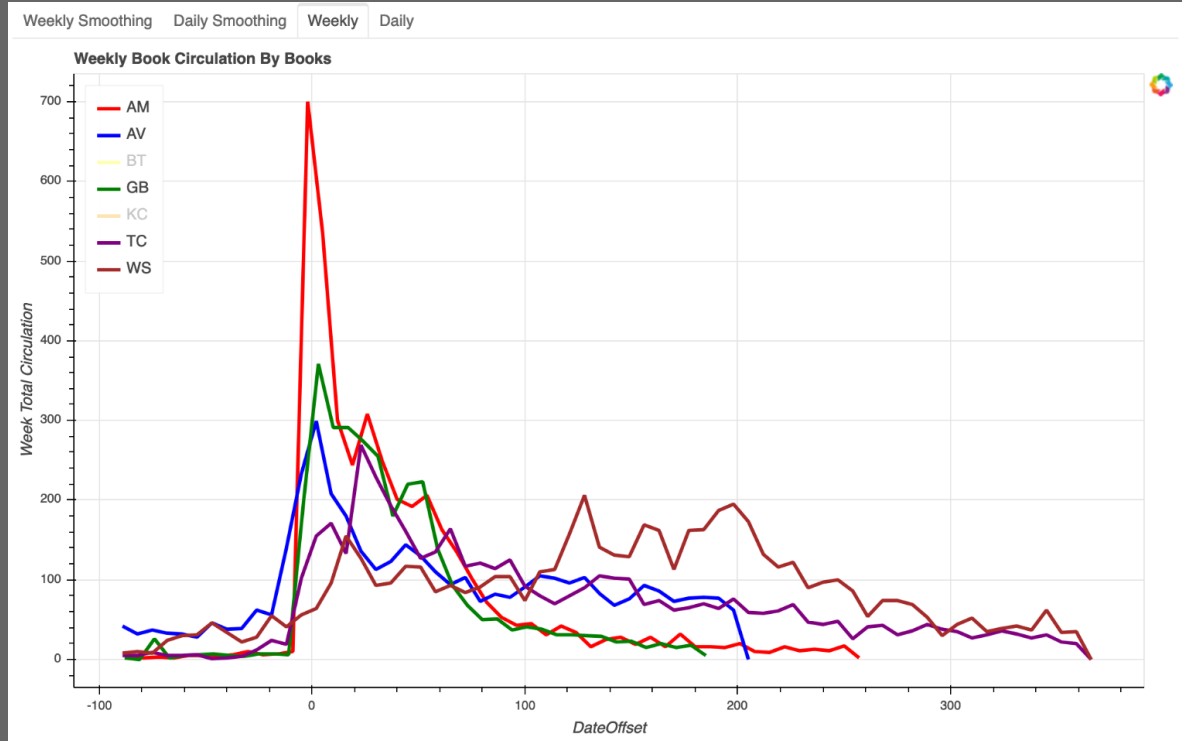
Our Catalyst Project Goals

- Create an interactive visualization dashboard for data associated with “One Book One Chicago” (OBOC)
- Demo the dashboard with CPL librarians and Lyris staff
- Distribute results at meetings and by publications

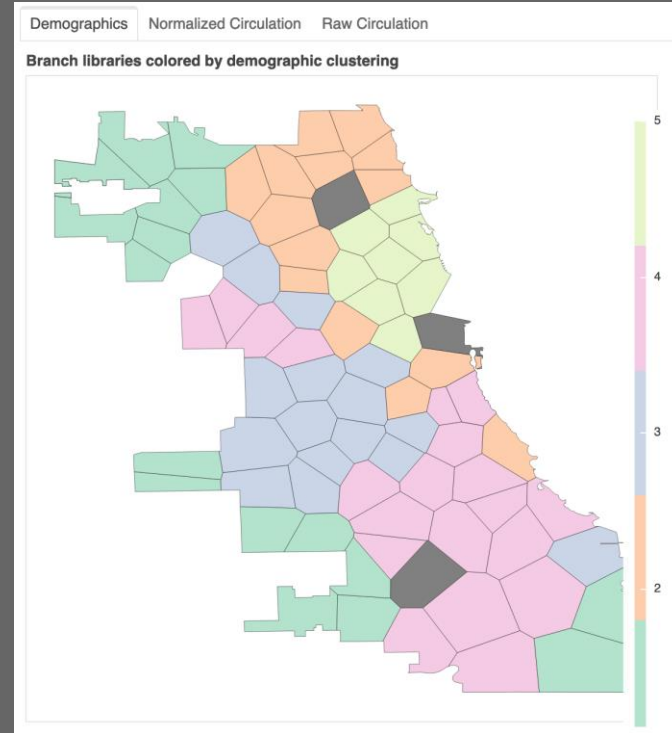
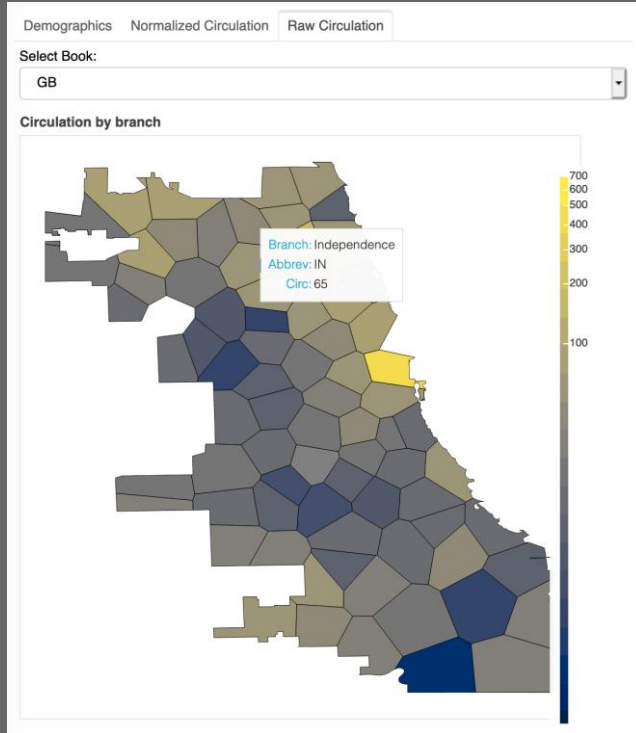
Timeline, Accomplishments, Challenges

- June 2018 - March 2020: pre-process and clean CPL and Census data; experiment with Bokeh for visualization; design and implement dashboard features
- Successes - we developed interactive features for librarians and other researchers
- Challenges - how to expand from one set of circulation records (CPL's OBOC) to a larger set of circulation records? What access to other data for modeling?

Dashboard Visualizations 1: Book time series



Dashboard Visualizations 2: Geographic



Lessons Learned

- OBOC seasons were not equal length; OBOC books are still under copyright. These require special handling. Modeling is difficult.
- The ephemerality of the OBOC program is a challenge. What is the archive, and how to capture and store it?
- The dynamic nature of OBOC data: CPL branches open/close, social media scraping is difficult; different media and multiple forms of participation (check out of print books, branch 'live' event attendance, e-book downloads, etc.)

Lessons Learned (continued)

- Combining circulation and event data can provide insights about the effect of promotional book events
- Our study can serve as a blueprint for assessments of large, library promotional events
- Results are particularly relevant to large public library systems as well as library consortia, but also useful for smaller library systems as well. Smaller library systems can compare their circulation pattern/s to a larger system.

Next Steps

- Move outward from our 2011-17 set to a larger sets of related books (first, the CPL recommendations for the OBOC books)
- Determine how to apply the model to more CPL circulation data and CPL programming events
- Determine how to make the dashboard useful to librarians in public library systems and in consortia

Next Steps (continued)

- Integrate OBOC promotional events into the dashboard
- Create visualization of text measures of the OBOC
- Integrate predictive modeling into the dashboard visualizations
- At present, the time-series shows all branches, but it is possible to make the dashboard work at the branch level
- Make our methods and models available to others (github, conferences, demos, etc.)
 - May require revisiting our CPL MOU

Questions and Comments?

To find out more, check the project website. Thanks!



Reading Chicago Reading

ABOUT RESULTS PEOPLE PARTNERS SPONSORS CONTACT

DEPAUL UNIVERSITY CHICAGO PUBLIC LIBRARY ONE BOOK ONE CHICAGO

About Reading Chicago Reading

Who reads? What do they read? How do they read? These are questions essential to the study of literacy, yet fine-grained answers to these questions are difficult to come by, as noted in *To Read or Not To Read*, a 2007 report from the NEA. Our project “Reading Chicago Reading” represents a rare opportunity to seek empirical answers to these questions within a large metropolitan area, with a wide variety of texts, and across a great diversity of readers.

“Reading Chicago Reading” is a digital humanities project that takes as its starting point the popular and much-imitated “One Book, One Chicago” (OBOC) program operated by the Chicago Public Library (CPL) since 2001. Each year (originally every six months), the library chooses one book around which to organize city-wide events, discussions, and other programming. For our purposes, the OBOC program acts as a natural experiment. The 25 different books that have been chosen since the program’s inception can be thought of as probes into the reading behavior of the city. The dependent variables of interest are, of course, the patterns of reading behavior that these probes elicit. We make use of scholarship in the history and sociology of reading as an interpretive baseline representing the standard approach. But our study, by bringing together Computer Science and Humanities research methods, promises new kinds of insight and new forms of quantifiable data.

Our research innovation is to combine data from texts, community demographics, circulation records, and social media that can yield predictions about who might be interested in a given book and why and that can help library staff visualize and anticipate the impact of book choices and event scheduling.

Search...

RECENT POSTS

Interactive maps from the ACH 2019 Conference August 5, 2019

What do CPL branch libraries have in common? February 12, 2019

Presentation at the Lyrasis Summit December 27, 2018

Presentations at DHCS 2018 November 23, 2018

Lyrasis Catalyst Fund Grant June 30, 2018

LATEST TWEETS

ReadingChiReading
@ReadChiReading

Thank you @LYRASIS for your support of our study of @chiphub's @1book1chicago program! Our

<https://dh.depaul.press/reading-chicago/>