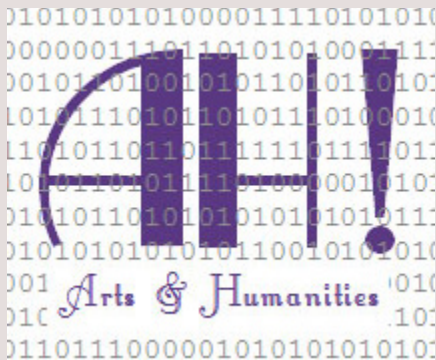# SIG AH Digital Newsletter

Association for Information Science & Technology: Arts & Humanities Special Interest Group

**2016-2017 Officers**
Chair: Jeremy McLaughlin (San Jose State University)
Chair-Elect & Social Media: Christian James (Catholic University of America)
Membership/Events: Julie Carmen (Eastern Washington University)
Communications: Tim Gorichanaz (Drexel University)
Website: Arjun Sabharwal (University of Toledo)
Archivist/Newsletter Editor: Marcus Ladd (Miami University)

## Message From The Chair

Where has 2016 gone?

It seems like only yesterday we were planning the 2016 Symposium and talking about the Annual Meeting in Copenhagen like a far-off idea. Here we are now in October and the AM is just days away. We wanted to end the previous year and start the new one (the ASIS&T fiscal year, that is) on a note of optimism: after a growth year in 2014-15 – earning SIG-of-the-Year – we maintained our momentum in 2016 and are working to grow the SIG and provide programs and services that promote our diverse interests within the Association.

We have several new faces joining the SIG AH leadership team and are currently planning several activities for the 2016-17 year. This includes webinars, the 2017 Virtual Symposium on Information and Technology in the Arts and Humanities, and (gasp) early ideas for the Annual Meeting next year outside Washington, D.C.

We look forward to seeing all of you at a future event. As always, if you have an idea for a speaker or a SIG AH webinar topic, or if you would like to contribute to the Newsletter or to the SIG please reach out to any of the Officers.

Finally – Save the Date! The 2017 Virtual Symposium on Information and Technology in the Arts and Humanities will be held April 18 and 19 2017, 12-4:30 pm EDT. The theme is The Information and the Technology of Open Humanities. More details to come.
We look forward to seeing you in Copenhagen and at a SIG AH event in the future.

Thanks,
Jeremy L. McLaughlin
*******************

Chair, SIG AH
Jeremy.mclaughlin@sjsu.edu

# Visualizations in Digital Collections: Recent Seminars
## Christian James, Catholic University of America

As a web librarian, I am frequently looking at research and case studies about digital collections. What are libraries doing with their digital collections? How can I keep my institution up to date on the latest technology and opportunities for innovation?

One new avenue for innovation may be visualizations. At two recent seminars, I explored the possibility of making on-demand graphic representations of text-based digital collections. At the April 2016 Images & Texts in Medical History in Bethesda, Md., Professor Ben Schmidt presented his web application, Bookworm, which takes plain text files and creates line chart graphs of keyword searches based on word frequency.

Many of us are already familiar with methods in the digital humanities. These include text mining large bodies of texts and representing them visually. For scholars, this usually means downloading, installing, and learning software like R or D3 (Data Driven Documents), which can be daunting. Applications like Bookworm allow anyone with a connected web browser to text mine or create visualizations for research or to just satisfy their own curiosity.
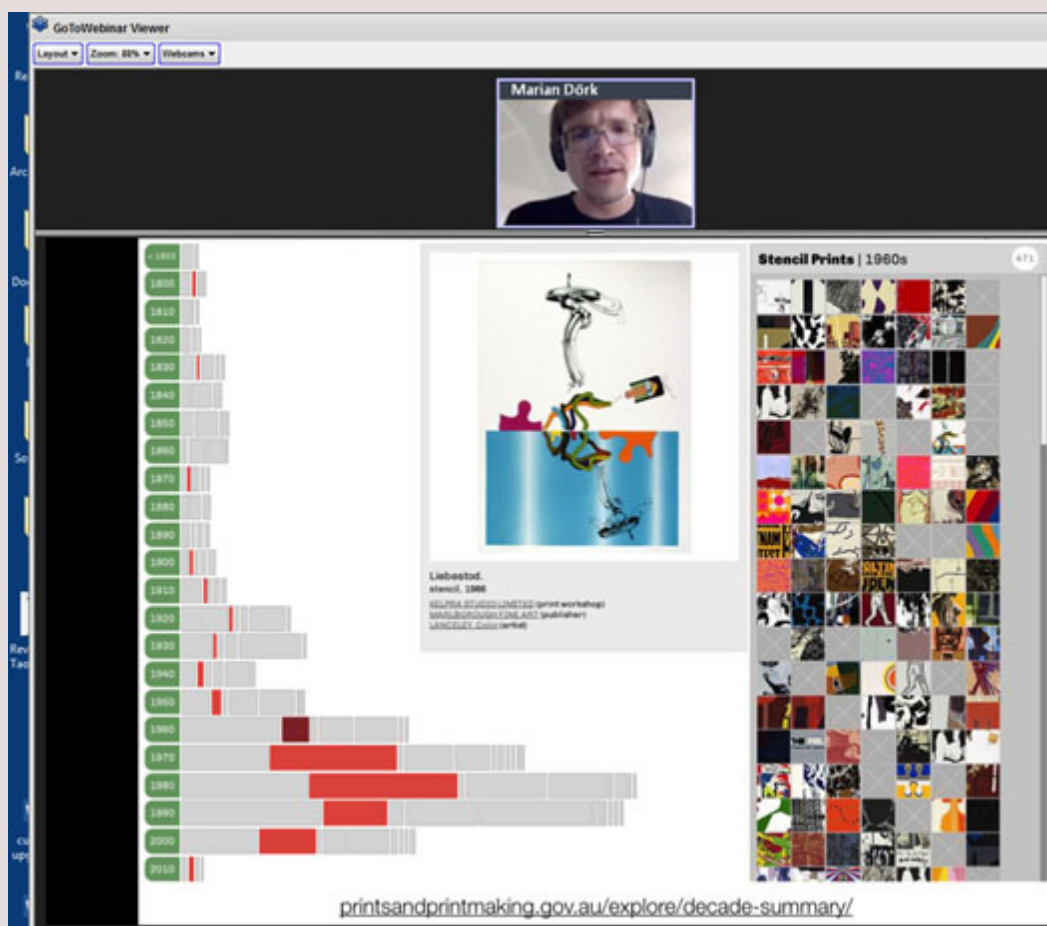


A recent ASIS&T webinar presented even more examples of digital collections that create on-demand visualizations. In July, ASIS&T SIG ED hosted Marian Dörk's presentation, "Visualizing Cultural Data: Experimental Interfaces for Digitized Collections." Professor Dörk presented experimental interfaces such as the German Digital Library, which asks "How does a collection of over 7 million cultural heritage objects look like?" The New York

Public Library (NYPL), like the German Digital Library, has created a massive collection of miniature document thumbnails to produce a stunning aesthetic. The National Gallery of Australia organizes its large array of thumbnails in an interactive timeline. This international perspective was welcome; while I was aware of NYPL's recent efforts, I was unfamiliar with many of the digital collections and initiatives from abroad.
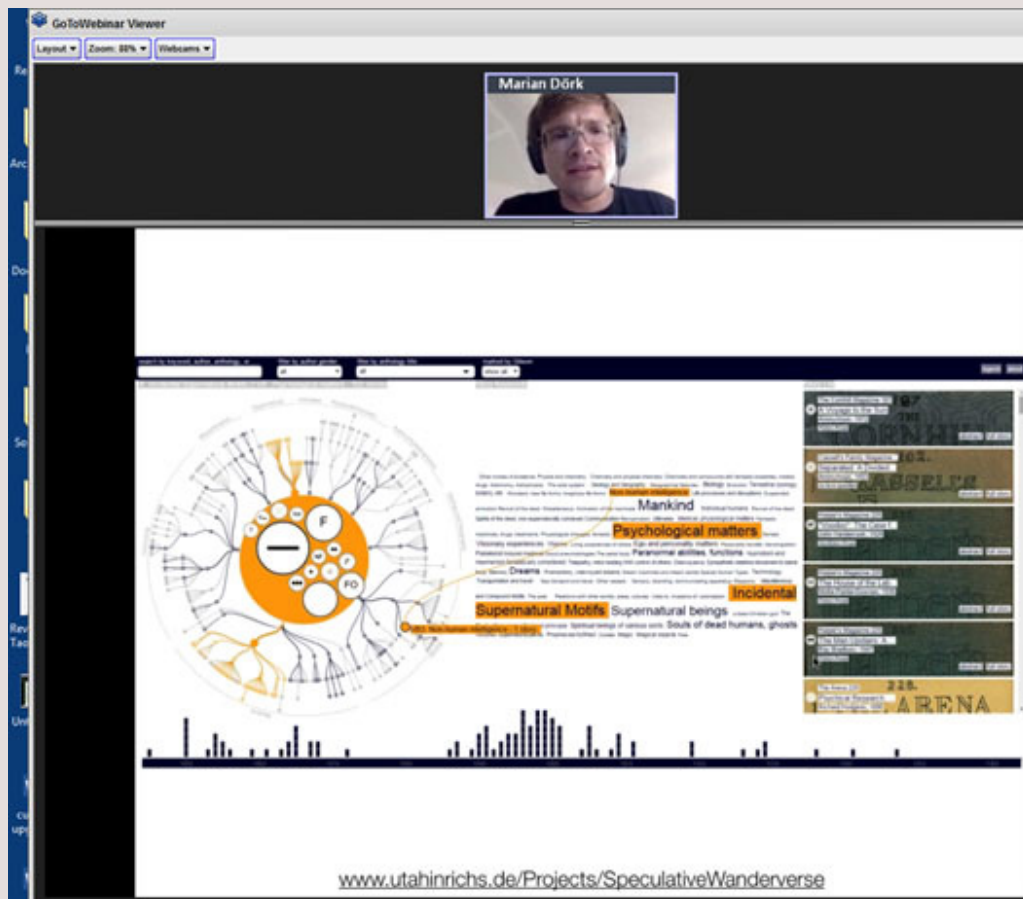


Perhaps more important than Dörk's examples were his theorizing of experimental interfaces. Experimental interfaces, Dörk says, can convey the extent of a collection, acknowledge a multitude of facets, show both details and abstractions, and provide significant context in ways that traditional interfaces cannot.

Dörk also argues that one subset of digital collection users are "flaneurs," curious explorers that enjoy critically spectating or reflecting from a distance, rather than pursuing a specific information need. Interactive interfaces would therefore be ideally suited to engaging this class of user.

These two presentations open many doors for future research and practice. How do users engage with these types of interfaces? Which ones are the most effective? How many (and what type of) libraries are implementing them, and why? These are questions that could help me and other librarians provide better access to information for the public.

# Research Update
## Marcus Ladd, Miami University of Ohio

For the past two years, I have been gathering analytics data of digital collections websites in Ohio. Although organizations like the Association of Research Libraries gather this data for their members, there are limitations to what they offer (i.e. it is only available to their members and only covers the statistics of research libraries). The idea for the project began in January 2015 at a meeting of the Ohio Digitization Interest Group, where the discussion turned to what "good" numbers would be for digital collections usage. Many of us had been gathering statistics and reporting them to our bosses, without any real metrics for evaluation. Because no one wanted to show their hand – again, not knowing how strong or weak their own numbers were – I created an anonymous survey to gather analytics data of digital collections websites.

I am currently completing my report of the survey's second year, this time with the help of Mandy Knapp at the State Library of Ohio (who knows a little more about statistics than I do). As we are still working within the Ohio DIG group, the actual sample size is not the largest (up to fifteen this year after ten last year), but we have responses from a variety of libraries and archives, including public, academic, corporate, and special. Areas we gather data on include the amount of labor put into developing digital collections; the size of the intended audience (e.g. student body population); the number of users, sessions, and page views; the average time spent on the digital collections website and the bounce rate (i.e. the percentage of visitors who leave more or less immediately without exploring the site); and the range of IP addresses that visitors come from (local, state, and national).
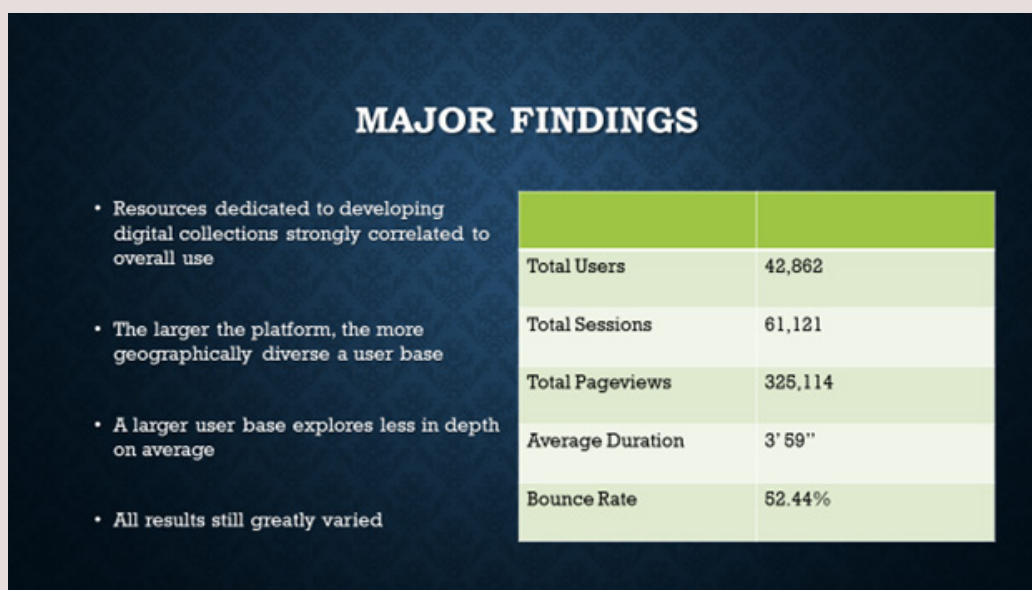


In addition to providing means and standard deviations for these areas to establish benchmarks (for, e.g., a medium-sized academic library to be able to compare its statistics to similar institutions), the report also includes correlations between the areas of assessment. Some of the revealed correlations are not surprising – that the larger the audience, the more likely they are to come from outside the local community – but the very strong correlation between hours spent on digital collections and overall usage provides the kind of specific data that is needed to justify continuing and expanding digitization efforts.

# 2016 Annual Meeting Sessions of Interest

Friday, October 14
9:00-17:00      Text Analytics and Sentiment Mining Using SAS 2 Day Seminar
8:30-12:30      Physical Samples, Digital Collections
13:30-17:30     Developing a Text Analytics Platform: From Taxonomies to Catonomies

Saturday, October 15
9:00-17:00      Text Analytics and Sentiment Mining Using SAS 2 Day Seminar
8:30-12:30      Visual Research Methods Workshop (SIG/VIS)
1:30-17:30      High Performance Sound Technologies for Access and Scholarship
17:30-20:00     miniBARcamp

Sunday, October 16
10:30           Student Design Competition Organizing Session
Noon            New Member/First Conference Brunch
5:30            Untangling Data Sharing and Reuse in Social Sciences

Monday, October 17
8:30            Diversity and Multiculturalism of LIS Education
8:30            A Right to Read for Machines
10:30           Plenary Session: Text and Data Mining Meets Pharmaceutical Industry
1:30            Digital Data Curation
3:30            How is Image Seeking and Use Studied (SIG VIS/AH/DL)
6:30            Presidents Reception Featuring Posters

Tuesday, October 18
8:30            Cultural Information Behavior
10:30           Preserving Intangible Heritage
10:30           Reading – New Methods, New Understanding
Noon            Annual Awards Luncheon
3:30            Learning and Information

The full program can be viewed at https://www.asist.org/events/annual-meeting/annual-meeting-2016/program/

## In Case You Missed It

The recorded webinars and presenter slides for the Second Virtual Symposium and our September 2016 webinars are now available.

2016 Virtual Symposium Proceedings

Technology and Publishing: The Work of Scholarship in the Age of its Digital Reproducibility

Linking Locally: A Technical Tour of the UNLV LOD Proj