
Data Visualization in Archival Finding Aids: A New Paradigm for Access

Anne Bahde

anne.bahde@oregonstate.edu
Oregon State University Libraries and Press
United States of America

Cole Crawford

cole.crawford@oregonstate.edu
Oregon State University, United States of America

Archival finding aids (also called collection guides) are meant to enhance access to archival collections for researchers, and have been presented online for almost two decades. However, researchers still struggle to understand and use them, and the poor functionality of finding aids can often impede the research process. Online finding aids frequently violate many tenets of basic web usability by presenting large blocks of text, complex collection hierarchies, and connections between relevant materials in arcane and unintuitive formats. Many scholars in the humanities have struggled with archival discovery and navigation systems in the course of their research or teaching; cumulatively, these individual annoyances add up to significant interference in the production of humanities scholarship in the humanities.

Ciaran Trace and Andrew Dillon have argued that user problems with archival finding aids may be rooted in the system of power inherent in these tools. According to Trace and Dillon, the traditional archival finding aid has always “reflected, privileged, enabled, and given control to the writer (archivist) more so than to the receiver (researcher)” (Dillon 2012). Users of online finding aids are merely receivers of the information, and in many systems have no control over how the information is presented, or even what information is presented. As a static document, the finding aid limits interpretive possibilities and aggregate analysis, because it cannot be (re)configured to meet users’ research needs. Passive, search-based finding aid systems hinder researchers’ potential for creativity, and obscure opportunities for serendipitous discovery.

Digital humanities modalities suggest compelling ways to disrupt this power paradigm. Johanna Drucker has called for “knowledge-generating visualizations,” which empower the user to produce new interpretations and arguments through manipulation and augmentation of the data presented (Drucker 2014). Tim Sherratt has advocated for the development of digital collection interfaces that enable users to visualize, augment, problematize, and critique collections and collections data (Sherratt 2011). These appeals invite questions about how users might respond to finding aids that present archival information in newly visual ways. What happens for the user when archival finding aids are stretched beyond traditional modalities to invite new interpretations of collections?

This presentation will introduce multiple models for including visualizations in both individual finding aids and discovery systems. Using case studies derived from pilot projects at Oregon State University’s Special Collections and Archives Research Center (SCARC), the presenters will discuss models that allow the researcher (and the archivist) to compare, match, recognize, distinguish, arrange, combine, construct, and organize data across a constellation of data points in ways that traditional, textual finding aids cannot. Models discussed will include proof-of-concept designs for interactive geographic timelines, force-directed network graphs, circle packing units, cluster force designs, alluvial diagrams, and treemaps that enable “distant reading” of a finding aid or corpus of finding aids. These visualizations allow users to identify patterns over time or space, relationships between collections, proportions between categories of data, and more. The presenters will explore how these visualizations can impact every stage of the research process, including topic exploration, identification of relevant archival collections, and establishing context and background.

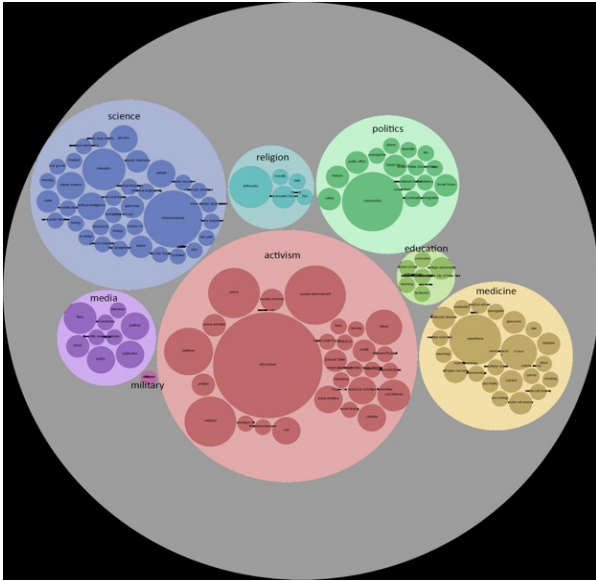


Figure 1: Example of circle-packing of letter keywords from “Global Bonds: Public Activism and Agency in the Letters of Linus Pauling”

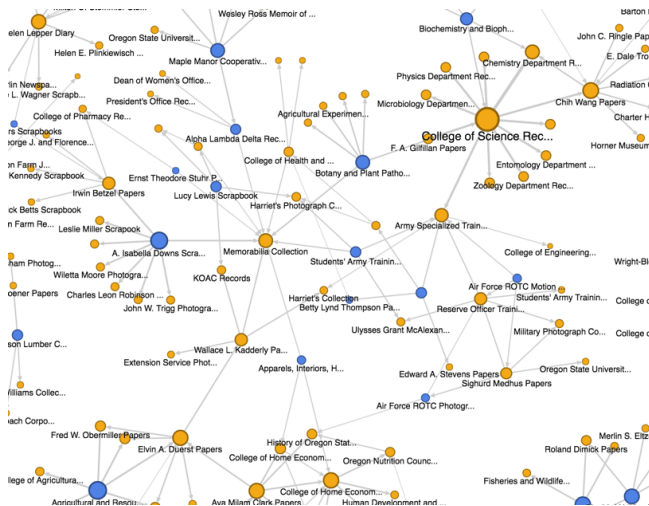


Figure 2: Example of force-directed network graph of related collections

Using several enhanced finding aids designed to promote access to archival materials, the presenters will examine practical considerations for researchers and archivists interested in creating visually enhanced digital finding aids. These considerations include platform selection, tool design, scalability, and data repurposing. SCARC faculty and staff work on multiple platforms, each tailored to specific project needs; choosing the right digital platform is critical to the long-term sustainability of a project, and this choice has only become more difficult with the proliferation of content management systems, digital publishing tools, and repository solutions.

Similarly, selecting the right development frameworks and types of visualizations for a certain data source can either enhance or limit a project’s potentialities. While the presenters have found success with libraries and tools such as d3.js, Leaflet, Sigma.js, and RAW, each archival discovery environment bears its own research complications which necessarily inform tool selection. The presenters will also talk about how both digital products and data production processes scale by considering project management and data entry workflows. Finally, the presenters will address how researchers can incorporate, clean, and repurpose existing data sources into new digital finding aids by examining a legacy finding aid that was repurposed and enhanced.

New research in the humanities often begins with consultation of finding aids and archival research environments. If access to archival collections is obscured by finding aid presentation and form, opportunities for progressive and provocative scholarship are concealed. Redesigned interfaces have the potential to invite new interpretations of collections, to encourage experimentation with research methods, and to plant new seeds of inspiration. Pairing new paradigms of visualization with recently proposed concepts of arrangement and description can open exciting possibilities; rhizomatic organization, for example, interrupts traditional hierarchy, structure, and power paradigms, and suggests a dynamic, acentric finding aid with multiple access points. If archivists create tools to make the archival research process adaptive and dynamic for users with different research needs, humanities scholars are empowered to pursue paths of inquiry never before revealed by traditional access tools.

Bibliography

- Drucker, J.** (2014). *Graphesis: Visual Forms of Knowledge Production*. Cambridge: Harvard University Press.
- Trace, C. and Dillon, A.** (2012). “The Evolution of the Finding Aid in the United States: From Physical to Digital Document Genre,” *Archival Science* 12 (4): 501-519.
- Sherratt, T.** (2011). It’s All About the Stuff: Collections, Interfaces, Power and People.” *Journal of Digital Humanities* 1 (Accessed October 15, 2106) <http://journalofdigitalhumanities.org/1-1/its-all-about-the-stuff-by-tim-sherratt/>