

Tools for the Digital Humanities:

A librarian's guide

Kristen Gallant, Elizabeth Lorang & Anneka Ramirez

School of Information Science & Learning Technology, University of Missouri

ISLT 9410: Emerging Technologies in Libraries

December 2014

Inside this guide:

- [Data preparation and early-stage research technologies](#)
- [Analysis technologies](#)
- [Publication technologies](#)
- [Additional Resources](#)

This brochure presents a collection of digital humanities tools that are currently being used by professional scholars, graduate students and persons in information science. The tools selected are aimed at humanities subject librarians who might be asked for ideas by their patrons. These tools are also an attractive option for librarians and information science professionals who are looking for options to digitize their institution's collections or who would like to publish research projects. Many of these tools are created with users in mind who may not have any IT experience. Because of this, more people have the capability of using technologies they wouldn't have before.

Data Preparation and Early-Stage Research Technologies

Scrivener - literatureandlatte.com/scrivener.php

Scrivener is a content generation tool for writers and researchers needing to structure long documents and projects. This tool allows users to build complex documents and provides a navigable menu to easily incorporate, compare, and structure information. Scrivener's website lists multiple projects for which it would be useful including thesis writing, screenplays, novels, and document translation. A free 30-day trial is available for download from the site and provides an in-depth tutorial for beginning users.

Nancy Bray (2013), through personal narrative, tells the story of why Scrivener works for her and how she used it to complete her final project for her master's degree in the article *Writing with Scrivener: A hopeful tale of disappearing tools, flatulence, and word processing redemption*.

Open Refine - openrefine.org

The data digital humanists want to analyze, remix, and publish are often messy: data sets might not use a controlled vocabulary, they might have redundancies, different formats might not work well together, and the data may include typographical errors, among other potential issues. OpenRefine is a tool for dealing with these issues and other messy data problems, and it allows users to clean and transform their data. The end result of using OpenRefine may not be as flashy as a dynamic visualization or slick web publication, but it does lay a crucial foundation for creating these other products as well as for supporting the long-term maintainability and usability of data.

For a look at how OpenRefine can be used within library and digital humanities environments, see Trevor Muñoz's (2013-2014) three-part series detailing the use of OpenRefine as part of the data curation activities for the New York Public Library's "What's on the Menu?" data: "[What IS on the menu? More work with NYPL's open data, Part One](#)"; "[Refining the Problem — More work with NYPL's open data, Part Two](#)"; and "[Borrow a Cup of Sugar? Or Your Data Analysis Tools? — More work with NYPL's open data, Part Three](#)."

Medieval Handwriting App -

play.google.com/store/apps/details?id=com.agbooth.handwriting.medieval

The Medieval Handwriting App was designed by Medieval scholars from the University of Leeds in West Yorkshire, U.K. It was designed to help both graduate and postgraduate students improve on their skills in paleography—"the art of analysing and reading handwriting" (Paleography 2011). The app allows users to try their hand at transcribing 26 different works that have been previously transcribed by scholars. When completed, the user can then compare their results with that of a professional. The app has options such as zooming in on particular words, the ability to save transcribed works, and sharing through file management software.

This free app can be used on Apple or Android devices. The app can be particularly useful for scholars who are not only challenged with just translating the language from Medieval texts, but also deciphering the letters themselves. While the app only offers a limited amount of manuscripts at this time, the user could practice their hand at paleography by reading manuscripts that have been made available online. A great place to use those skills is [Manuscripts Online](#) which not only has digitized manuscripts but also samples of free-hand writing from the 10th to the 15th centuries.

Analysis Technologies

Voyant - voyant-tools.org

Voyant is a free, web browser-based tool for text analysis. Users can input text for analysis via URL, copy and paste, or file upload, and Voyant can process a single file or a collection of files. Uploaded files can be in a variety of formats: doc or docx; html; xml; rtf; and pdf. Once a user provides content to Voyant, the tool opens up an analysis pane, which offers graphical visualizations and textual descriptions of the file or corpus, including "number of words, number of unique words, longest documents, highest vocabulary density, most frequent words, notable peaks in frequency, and distinctive words" (Sinclair & Rockwell, 2014).

Voyant is more sophisticated than other web-based text analysis tools, such as Wordle, and easier to use than programming-intensive software (or writing code from scratch). Voyant's features make it a good middle-weight text analysis tool, and it is well-suited for both teaching and research environments. A user manual and screencasts [are available](#), and the project has also created [some excellent documentation](#) outlining its principles and a variety of use cases.

Piktochart - piktochart.com

Piktochart is an easy-to-use infographic creation website that provides data visualization and presentation tools, which many beginning digital humanists will find accessible. It is also appealing to educators who want to represent subjects, research, and data in a colorful and easy-to-read format for classrooms, conferences, and seminars. Accounts are free to create and provide a number of customizable templates and charts to begin; however, users can pay to upgrade to a pro account and gain access to a wider range of tools.

While there are many ways to represent data, Piktochart provides tools to help arrange information and data to tell a story. Michelle Moravec, Associate Professor of History at Rosemont College, created a Piktochart infographic, titled [digitally](#), about introducing digital humanities to classrooms of higher education institutions. And, [artwis.com](#) uses Piktochart to create infographics to display art history information, like in this [Chinese Ceramics](#) blog post.

Gephi - gephi.github.io/

Gephi is an open-source platform for analyzing networks and other systems. According to the Gephi website, the software aims to help researchers make hypotheses and discover patterns. Humanities researchers have used Gephi for social network and citation analyses. For an example of a historian using Gephi, see Brian Sarnacki's (2014) "[Playing with Gephi](#)" blog post and his work-in-progress on Grand Rapids, Michigan [social networks](#), as well as Amanda Visconti's (2012) "[View DHQ: Getting Started with Gephi for DH Vis.](#)"

Gephi is still in beta, and unlike some other resources in this guide, it can be intimidating to the beginning user. Librarians may want to consider offering a Gephi workshop or series of workshops on campus, to help researchers become familiar with the software's file formats, interface, and potential. Because Gephi is used by scientists, journalists, and humanists, there may already be an individual or group on your campus with the knowledge necessary to lead such an event.

Palladio - palladio.designhumanities.org

Developed by the Humanities + Design research lab at Stanford University, Palladio is another web platform for creating visualizations of data. Unlike Gephi, Palladio has been developed specifically for a humanities audience, and its tagline is "Humanities thinking about data visualization." In particular, Palladio is useful for visualizing and discovering multi-dimensional relationships in time and space. To create a Palladio project, users upload their data in a spreadsheet or CSV file or in the form of a query in the SPARQL format (for more advanced users). Once a user has created a Palladio project, she can save the Palladio project locally (no data is stored on Palladio servers). The user can then distribute and share that Palladio project or re-upload it to Palladio for continued development.

A work-in-progress and an active research project, Palladio does not have the most intuitive of interfaces, and the documentation could be more robust. Fortunately, Miriam Posner (2014) has written a [very useful blog post](#) with a detailed how-to. The post also outlines situations in which researchers may or may not want to use Palladio, and the post also includes a detailed how-to. In addition, prospective users may want to check out Mark Braude's (2014) [HASTAC post](#) on a variety of projects that have experimented with Palladio.

Publishing Technologies

Scalar - scalar.usc.edu/

Scalar is a web publishing platform that is designed for people with little to no technical experience. Scalar is free and open-source and places strong emphasis on multimedia publication. This allows for any number of images, video and audio to be published alongside accompanying text. "Scalar is based on the semantic web standard RDF, and leverages widely accepted ontologies like Dublin Core, SIOC, and ArtSTOR" ("About Scalar," 2014). Scalar is an ideal platform for the publication of born digital objects and can be used to create web pages that read like digital books or articles that have linkable content.

Some projects that are currently using Scalar are: [Survivor's on Schindler's List](#) and [Cesar Chavez's Video Collection](#). Both of these projects are about completely different subjects, but both show how Scalar is great for combining different forms of media in one venue.

Camtasia - techsmith.com/camtasia.html

Camtasia is a video editing tool that works on Mac and Windows platforms and creates videos viewable on multiple devices. This tool is beneficial for supplementing lecturing, creating videos to document research and projects, producing tutorials for a variety of environments and purposes. Camtasia offers screen capture, audio effects, transitional frames, and a media library to import images, audio clips, and video. Multiple tracks allow users to easily create layers of editing for a seamless finished work.

Camtasia videos are easily uploaded to Youtube; for example Youtube user Rebecca Jordan uploaded the video [Camtasia Screencast Literary Response to "The Road Not Taken" By: Robert Frost](#). Kathleen Carlson's article [Delivering information to students 24/7](#), provides an examination of the process to select and then implement Camtasia for library usage tutorials. For a look at a finished Camtasia screencast project, check out the Enoch Pratt Free Library's tutorial for use of their [Literature Resource Center](#).

Omeka - omeka.org

Omeka is an "open source web-publishing platform" that is commonly used for locally hosted digital asset management (DAM). Because Omeka is free DAM software, it is often used by institutions, repositories or organizations working on small-scale budgets. Omeka is designed for people with little to no html coding or IT experience. The user-friendly software allows for the creation, preservation and sharing of digital assets such as scholarly works, images, video and audio files. Omeka encodes all of its record metadata using Dublin Core, which is a flexible schema that works across multiple platforms. Omeka is used by libraries, museums, archives and organizations to create an environment that allows various forms of digital content to be accessed.

Omeka also offers a hosted option-Omeka.net for those institutions that cannot or do not want to host their own website. Some Omeka-based projects that you might want look at are: the [Digital Public Library of America](#), the [History Harvest](#) and [100 Years: The Rockefeller Foundation](#).

Additional Resources

Digital History Project. <http://digitalhistory.unl.edu/>

A Digital Toolbox for Historians <http://www.pinterest.com/ahahistorians/a-digital-tool-box-for-historians/>

DiRT: Digital Research Tools <http://dirtdirectory.org/>

A Guide to Digital Humanities. <http://sites.northwestern.edu/guidetodh/>

Text Analysis Portal for Research (TAPoR) <http://www.tapor.ca/>

References

100 Years: The Rockefeller Foundation. (2014). Retrieved from <http://rockefeller100.org/>

About Scalar. (2014, December 3). Retrieved from <http://scalar.usc.edu/features/web-standards/>

Barrett-Wilt, K. (2013 September 5). Chinese Ceramics [sic], an infographic [Web log post]. Retrieved from <http://www.artwis.com/articles/chinese-ceramics-an-infographic/>

Bray, N. (2013). Writing with Scrivener: A hopeful tale of disappearing tools, flatulence, and word processing redemption. *Computers And Composition: An International Journal For Teachers Of Writing*, 30(3), 197-210. doi:10.1016/j.compcom.2013.07.002

Carlson, K. (2009). Delivering Information to Students 24/7 with Camtasia. *Information Technology & Libraries*, 28(3), 154-156.

Digital Public Library of America. (2014). Retrieved from <http://dp.la/>

Enoch Pratt Free Library. How to use the Literature Resource Center [Video tutorial]. Retrieved from <https://www.prattlibrary.org/special/camtasia.aspx?id=1762>

Heddaya, M. (2014, November 7). Finally, an app for transcribing medieval manuscripts. *Hyperallergic*. Retrieved from <http://hyperallergic.com/161301/finally-an-app-for-transcribing-medieval-manuscripts/>

Jordan, R. (2013 April 18). Camtasia Screencast Literary Response to “The Road Not Taken” By: Robert Frost [Youtube video]. Retrieved from <https://www.youtube.com/watch?v=xRRustUJwxs>

Moravec, M., Ph.D. (2014). *digitally* [Infographic]. Retrieved from <https://magic.piktochart.com/embed/2823892-digiped>

Munoz, T. (2013, August 8). What IS on the menu? More work with NYPL's open data, part one [Web log post]. Retrieved from <http://www.trevormunoz.com/notebook/2013/08/08/what-is-on-the-menu-more-work-with-nypl-open-data-part-one.html>

Munoz, T. (2013, August 19). Refining the problem — More work with NYPL's open data, part two [Web log post]. Retrieved from <http://trevormunoz.com/notebook/2013/08/19/refining-the-problem-more-work-with-nypl-open-data-part-two.html>

Munoz, T. (2014, January 10). Borrow a cup of sugar? Or your data analysis tools? — More work with NYPL's open data, part three [Web log post]. Retrieved from <http://trevormunoz.com/notebook/2014/01/10/borrowing-data-science-tools-more-work-with-nypl-open-data-part-three.html>

Posner, M. (2014, November 23). Getting started with Palladio [Web log post]. Retrieved from <http://miriamposner.com/blog/getting-started-with-palladio/>

Sarnacki, B. (2014, January 9). Playing with Gephi [Web log post]. Retrieved from <http://www.briansarnacki.com/gephi/>

Sarnacki, B. (n.d.). GR social networks. Retrieved from <http://www.briansarnacki.com/smallcity/networks/index#>

Sinclair, S., & Rockwell, G. (2009). Voyeur tools: See through your texts. Retrieved from <http://hermeneuti.ca/voyeur>

Sinclair, S., & Rockwell, G. (2014). Principles of Voyant tools. Retrieved from <http://docs.voyant-tools.org/context/principles/>

Tillotson, D. (2011, December 2). What is paleography? *Medieval Writing*. Retrieved from <http://medievalwriting.50megs.com/whatis.htm>

Title and heading images. Digital Bokeh. Created by Heather Katsoulls. Image recolored. Retrieved from <http://goo.gl/1sPVdE>

Visconti, A. (2012, July 11). View DHQ: Getting started with Gephi for DH vis (part II of II) [Web log post]. Retrieved from <http://www.literaturegeek.com/wp-content/uploads/2012/07/viewdhqpt2>