Digital Humanities for the History of the Book

UCLA California Rare Book School, August 2021

Instructor: Dr. Catherine DeRose
Email: catherinederose@gmail.com

Class materials
GitHub Link

Zoom link
Passcode (see welcome email)

Course Description

This course will introduce digital humanities methods and workflows by applying them to questions from literary studies and history. We will begin by examining the processes by which written and printed materials become digital, asking what is gained and lost along the way. What kinds of claims can we make about the "data" as a result of how it was curated or transformed? What questions can we ask when we increase our scale of study to hundreds or thousands of texts? Over the week, we'll work closely with texts and their accompanying metadata to see what new information we might glean by incorporating digital tools into our practice.

Digital humanities is an expansive, rapidly developing area of research. This course is designed to give you a foundation, introducing you to approaches and resources you can continue to draw on after the week concludes. You will gain hands-on experience working with popular open source tools used by digital humanities practitioners. You will also learn best practices for developing your own digital humanities projects, from curating a dataset and identifying a suitable method of analysis to creating effective and compelling data visualizations you can share.

This course is intended for students, faculty, and library staff who are interested in taking on or supporting digital humanities projects, or who are just looking for an introduction to the field. No prior programming experience or project is required. Sample data will be provided.

Since this course will be taught remotely, we will have a combination of synchronous and asynchronous activities. The synchronous portion of each day will be divided into three components: discussion (where we'll consider case studies and methods at a conceptual level), workshop (where we'll carry out a method using a digital tool), and open lab (where you'll have the opportunity to develop your own projects or continue experimenting with the sample data). The asynchronous portion will consist of readings that will help inform our discussions, project reviews that will provide an opportunity for engaging with digital humanities work more closely, and hands-on practice exercises that will reinforce the workshop material.

Learning Objectives

Through this course, you will:

- Develop an understanding of data—its history, contexts, and uses—in the (digital) humanities
- Learn how to prepare datasets of varying scales
- Identify appropriate methods for pursuing different research questions
- Gain experience working with out-of-the-box data analysis and visualization software
- Establish a foundation for continuing with research in the digital humanities.

Readings & Software

All texts for this class are publicly and freely available. In addition to the readings we'll discuss in class, there will also be optional readings in our course GitHub repository that you can follow-up with if you become interested in an area and want to do a deeper dive. **In preparation for the first class, please read: "Big? Smart? Clean? Messy? Data in the Humanities," by Christof Schöch.**

The software we're using in this course is either open source or has a free account option that we'll use. For software that requires installation, please try to have it installed prior to the class in which we'll be using it. I'll be available over Zoom to assist during the open lab sessions with any installation hiccups. Participants may use Windows or Mac computers.

Project Reviews

To help concretize the material we'll be covering in discussion, I've selected a project to accompany the reading for most days. When interacting with these projects, I encourage you to consider:

- the visualization
    - What is the visualization trying to communicate—is it making an argument or is it more exploratory?
    - In what ways is it effective or ineffective?
- the data
    - Is it open or in copyright? How can you tell?
    - How was it collected, processed, or transformed?
    - What's missing from the data?
- the method
    - What questions can—or can't—you ask of the data as a result of the technical approach?
- the team
    - Who are the creators behind the project?
    - What are their areas of expertise?
- the results
    - What are your takeaways from the project?

In addition to providing sample use cases for methods we'll be covering, these projects will also offer insight into how digital humanities projects come together (often through collaboration) and find expression online.

Schedule (Pacific Time)

August 2 — Data in the Humanities

9:00am — 10:30am	Discussion: Digital Humanities Participant and course introductions Overview — what is digital humanities? what do we mean by "data"? Discussion of the reading

11:00am — 12:30pm	Workshop: OpenRefine
2:00pm — 3:00pm	Open Lab
Your time	Reading (please have read before next class)
	"Feminist Data Visualization"
	Practice exercise – link will be shared on GitHub

August 3 — Data Visualization in the Humanities

9:00am — 10:30am	Discussion: Data Visualization
	Overview — how do we design with data and humans in mind? Discussion of the reading Sample visualizations
11:00am — 12:30pm	Workshop: Tableau Public
2:00pm — 3:00pm	Open Lab
Your time	Reading & Project (please look through before next class)
	"Seven Ways Humanists are Using Computers to Understand Text"
	Robots Reading Vogue (N-Gram and Topic Modeling)
	Practice exercise – link will be shared on GitHub

August 4 — Texts

9:00am — 10:30am	Discussion: Text Analysis
	Overview — what is it? what questions and data is it well suited for? key terms? Discussion of the reading Sample projects
11:00am — 12:30pm	Workshop: Named Entity Recognition with Stanza/Google Colab
2:00pm — 3:00pm	Open Lab
Your time	Reading & Project (please look through before next class)
	"Network Theory, Plot Analysis"
	Six Degrees of Francis Bacon
	Practice exercise – link will be shared on GitHub

August 5 — Networks

9:00am — 10:30am	Discussion: Network Analysis

	Overview — what is it? what questions and data is it well suited for? key terms? Discussion of the reading Sample projects
11:00am — 12:30pm	Workshop: Gephi
2:00pm — 3:00pm	Open Lab
Your time	Reading & Project (please look through before next class) "The True Colors of America's Political Spectrum Are Gray and Green" The Atlas of Early Printing Practice exercise – link will be shared on GitHub

- -

August 6 —Maps & Publishing

9:00am — 10:30am	Discussion: Geospatial Analysis & Sharing Digital Humanities Work Overview — what is it? what questions and data is it well suited for? key terms? Discussion of the reading Sample projects
11:00am — 12:30pm	Workshop: ArcGIS & StoryMaps
2:00pm — 3:00pm	Open Lab
Your time	Practice exercise – link will be shared on GitHub