Texts, Maps, and Networks: Introduction to Digital Humanities

2018 California Rare Book School

Catherine DeRose
catherine.derose@yale.edu

Class slides
[URL to class site removed]

Overview

This week-long course offers an introduction to the field of digital humanities, broadly considered. We will begin with an overview of the field and a general discussion of data in the humanities, looking at ways to collect, clean, and visualize it. The remaining days of the course will consist of deep dives into specific areas of digital humanities, including text mining, network analysis, and mapping. We will conclude the week by discussing options for sharing your work online.

Participants in the course will gain hands-on experience working with out-of-the-box, open source tools. Sample datasets will be provided, though there will also be opportunities for participants to build their own. Short, open access readings for each day will be available online. Throughout the course, we will move between aggregate, distant views of texts and traditional close readings to consider what is gained and lost in the transition.

Schedule

Data in the humanities – Monday, July 30

Introductions and course overview 9:00am – 10:30am

- Participant and course introductions
- What is digital humanities?
- What is data in the humanities?
- Working with online, messy data

- break -

Data collection and cleaning 11:00am - 12:30pm

- Creating your own dataset by hand
- Finding curated datasets online
- Web scraping with the Unix shell
- Cleaning your data with OpenRefine and Lexos

- lunch -

- CalRBS Group Lunch, GSEIS Salon, 2nd floor

Data visualization fundamentals 1:30pm - 3:00pm

- Data visualization best practices
- Creating graphs in Tableau Public

- break -

User experience and assessment 3:30pm - 4:30pm

- Formal and informal project assessment
- User-driven digital humanities

Installation and lab time 4:30pm - 5:00pm

Texts – Tuesday, July 31

What is computational text analysis? 9:00am - 10:30am

- Overview — what is it? what questions and data is it well suited for? key terms?
- Discussion of reading
- Sample projects

- break -

Out-of-the box tool exploration 11:00am - 12:30pm

- Voyant
- AntConc
- DataBasic.io

- lunch -

Hands-on experiments with text mining 1:30pm - 3:00pm

- Topic modeling with the Topic Modeling Tool
- Named entity recognition with Stanford's NER GUI

- break -

Lab time 3:30pm - 5:00pm

CalRBS Mixer 5:00pm

- UCLA YRL Special Collections

Networks – Wednesday, August 1

What is network analysis? 9:00am - 10:30am

- Overview — what is it? what questions and data is it well suited for? key terms?
- Discussion of reading
- Sample projects

- break -

Network data 11:00am - 12:30pm

- Assembling node and edge lists by hand
- Generating edge lists with Python

- lunch -

Hands-on experiments with networks 1:30pm - 3:00pm

- Creating graphs in Gephi

- break -

Lab time 3:30pm - 5:00pm

CalRBS Lecture 5:30pm

- GSEIS 111

6:30pm Reception

- GSEIS Salon, 2nd floor

Maps – Thursday, August 2

What is geospatial analysis? 9:00am - 10:30am

- Overview — what is it? what questions and data is it well suited for? key terms?
- Discussion of the reading
- Sample projects

- break -

Geospatial data 11:00am - 12:30pm

- Geocoding by hand
- Geocoding with Geonames API

- lunch -

Hands-on experiments with maps 1:30pm - 3:00pm

- Creating a map in AcrGIS online
- Presenting your map in Esri's Story Maps
- Creating a map in Carto

- break -

Lab time 3:30pm - 5:00pm

Sharing your work online – Friday, August 3

Tour of the Scholarly Innovation Lab 8:55am - 10:30am

- Meet in the entrance of the Charles E. Young Research Library

- break -

Web hosting 11:30am - 12:30pm

- Building a website with GitHub pages

- lunch -

Next steps and course evaluation 2:00pm - 3:30pm

- Follow-up tutorials and workshops
- CalRBS evaluation