

Data in the Humanities

Wk. 01: Introductions; Platforms; AI



Today's plan

1. 🙌 Introductions
2. 🌍 Getting our bearings: the class, platforms, questions, tools
3. 📖 Discuss readings
4. 🛠️ Overview of some training resources, funding opportunities, campus infrastructure

Introductions 🙌

- Your name & department?
- What do you work on / what are you interested in?
- What brings you to this class / what do you hope to learn?

Getting our bearings 🌍

The Center for Digital Humanities

CDH projects



Getting our bearings 🌐

Platforms we'll use

<https://gwijthoff.github.io/CDH507/>



Getting our bearings 🌐

Main questions

1. What is DH, past, present, and future?
2. How do quantitative methods operate in the humanities?
3. Why are quantitative, computational, and data-driven methods experiencing such a resurgence now?
4. How can we distinguish a) genuinely new arguments and discoveries produced by those methods from b) techno-solutionism?

Getting our bearings 🌐

Tools introduced

- command line `>_`
- tabular data `author,title,year`
- OpenRefine `facetCount(value) > 1`
- Jupyter notebooks `import pandas as pd`
- web scraping `from bs4 import BeautifulSoup`
- GIS `[-84.375, 31.952162238024975]`

VisiData for EDA

OPERATOR	count	percent	histogram
UNKNOWN	22376	31.42	*****
SOUTHWEST AIRLINES	7752	10.55	*****
BUSINESS	5868	7.99	*****
AMERICAN AIRLINES	4337	5.98	*****
DELTA AIR LINES	2917	3.94	*****
FEDEX EXPRESS	2789	3.69	*****
UNITED AIRLINES	2194	2.99	****
US AIRWAYS	1885	2.57	****
UPS AIRLINES	1773	2.41	***
SKYWEST AIRLINES	1769	2.41	***
JETBLUE AIRWAYS	1740	2.37	***
EXPRESSJET AIRLINES	1347	1.83	**
AMERICAN EAGLE AIR	1041	1.42	**
ENVOY AIR	883	1.20	*
ALASKA AIRLINES	865	1.14	*
REPUBLIC AIRLINES	804	1.09	*
NESA AIRLINES	693	0.94	*
AIR WISCONSIN AIRL	623	0.85	*
PSA AIRLINES	577	0.79	*
PRIVATELY OWNED	516	0.70	*
PHI INC	491	0.67	*

PUL Makerspace



Selene photometric scanner



BEARINGS

- Wk. 01 | Jan 31 | Introductions; Platforms; AI

THEORIES

- Wk. 02 | Feb 07 | Histories of the Humanities
- Wk. 03 | Feb 14 | Histories of DH: From IBM Up to the Big Tent
- Wk. 04 | Feb 21 | DH Today: Meaning & Signal

- Wk. 05 | Feb 28 | Is “DH” Now Just “H”?
- Wk. 06 | Mar 06 | Quantitative Methods: Models, Generalization, & Reproducibility

METHODS

- Wk. 07 | Mar 20 | Scraping & Curating Structured Data
- Wk. 08 | Mar 27 | Maps & Spatial Humanities
- Wk. 09 | Apr 03 | Networks & Artifacts
- Wk. 10 | Apr 10 | Filtering and Modeling Data
- Wk. 11 | Apr 17 | Refusal

PROJECTS

- Wk. 12 | Apr 24 | Project Proposal Presentations

Wk. 1 Readings

- Meredith Whittaker, “The Steep Cost of Capture,” *Interactions* 28, no. 6 (November 10, 2021): 50–55.
- Matthew Kirschenbaum, “Prepare for the Textpocalypse,” *The Atlantic*, March 8, 2023.
- James Somers, “A Coder Considers the Waning Days of the Craft,” *The New Yorker*, November 20, 2023.
- Adam Rogers, “Our Greatest AI Visionary Isn't Sam Altman or Bill Gates. It's Ted Underwood,” *Business Insider*, December 2023.

training resources, funding opportunities, campus infrastructure

Resources

Resources

campus training

- CDH [Graduate Training Grants](#)
- Spring 2024 [PUL workshops](#) related to "data & computation."
- Research Computing's [Training for the Humanities & Social Sciences](#) page
- PRDS [Research Data Stewardship Program](#)

general resources

- Many linked on our course site's [Resources page](#)
- [dh-notes](#), a crowdsourced list of DH conferences, workshops, and funding sources
- [DHRIFT](#) (Digital Humanities Research Institute)