

English 58I: Digital Tools and Critical Theory

Ted Underwood

M 1:00 - 2:50

office hours: 1:30 - 2:30, 3:30 - 4:00, and by appt.

Spring 2012

office: EB 339

“Digital humanities” is a broad rubric, which includes new objects of study (video games, electronic literature) as well as proposals for reforming academic communication (the open access movement).

While we’ll discuss that whole panorama, I’d like to focus this course on the relationship between information technology and critical theory — because I see developments there that could be immediately useful for scholars studying a wide range of disciplines and periods. We’ll ask how humanists’ interpretive strategies may already have been shaped by technology, and explore ways of using technology to enlarge the range of strategies we have available.

Texts.

Matthew K Gold. ed. *Debates in the Digital Humanities*. 2012.

Franco Moretti. *Graphs, Maps, Trees*. 2005.

Stephen Ramsay. *Reading Machines: Toward an Algorithmic Criticism*. 2011.

Stefan Gries. *Quantitative Corpus Linguistics with R*. 2009.

Things we’ll be doing.

I have established a discussion board on a Moodle website for the course (<https://courses.las.illinois.edu/my/index.php>) and I hope discussion will often start there and then continue into class. Don’t feel that you have to write a comment before every class; about half of the time is fine. Questions, brief manifestoes, cries of “eureka” are all welcome.

I’ll propose that we create a group blog in Wordpress for more public presentation of work-in-progress.

There will be a number of short coding assignments, which you can keep doing until they work. There will also be a lot of “praxis” assignments that basically encourage you to play around creatively with tools that map discourse in variety of different ways.

The final project for the class will be an essay, somewhat on the short side because we’re doing a lot here besides writing. Aim at 10-12 pages. The topic is wide open, and does not

have to look digital or quantitative at all. I would tend to recommend writing a brief essay on literary or intellectual history in a period that interests you, *subtly* informed by new ways of mapping an archive or analyzing text. “Subtly” here could mean simply that the idea occurred to you as a consequence of work related to the course. Naturally, you are also free to write a paper that reflects explicitly on some theoretical question we explore. But I’d like the project to genuinely advance your other interests; the central agenda of this course is to show that digital approaches can be integrated unobtrusively into critical projects humanists are already pursuing. (In fact they already have been integrated ...)

A lot of the examples we’ll be considering are in English, and are located somewhere between 1600 and 1923. If you’re interested in a different language and/or period, see me at some point before March, so I can help you find resources that will be more directly relevant to your interests. It should be possible.

January 23rd. What are “digital humanities,” and are we already practicing them?

- Matthew Kirschenbaum, “What Is Digital Humanities and What’s It Doing in English Departments.” 2010, rpt. in *Debates*, pp. 3-11. If you can’t get *Debates* yet, the original version is available at: <http://mkirschenbaum.files.wordpress.com/2011/03/ade-final.pdf>
- Peter Stallybrass, “Against Thinking,” *PMLA* 122.5 (2007): 1580-85.
- Johanna Drucker, “Humanistic Theory and Digital Scholarship,” in *Debates*, pp. 85-95.
- Ted Underwood, “The methodologies we argue about, and the ones we use quietly.” MLA 2012, Seattle. <http://www.authorstream.com/Presentation/tedunderwood-1309452-mla2012final/>

January 30th. Relatively distant.

Theory: • Franco Moretti, *Graphs, Maps, Trees*, 1-33.

- Jean-Baptiste Michel*, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, William Brockman, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden*. “Quantitative Analysis of Culture Using Millions of Digitized Books.” *Science*, 12/16/2010. (Moodle)
- Ben Schmidt, “Age Cohort and Vocabulary Use,” *Sapping Attention*, 2011. <http://sappingattention.blogspot.com/2011/04/age-cohort-and-vocabulary-use.html>

Praxis: Go to the Google ngram viewer (<http://books.google.com/ngrams/info>) and explore whatever and however you like. Then compare Bookworm (bookworm.culturomics.org). Finally check out the SEASR Correlation Miner (<http://leovip026.ncsa.uiuc.edu/Correlation/>) Find a correlation that seems interesting or surprising. Do you think it has any real significance?

Critique of praxis: Matt Jockers, “Unigrams and bigrams and trigrams, oh my,” <http://www.stanford.edu/~mjockers/cgi-bin/drupal/node/53>

February 6th. Relatively close.

Theory: • Stephen Ramsay, *Reading Machines*, “Preconditions” and Chapter 1: “An Algorithmic Criticism,” pp. ix - 17. Also “Postconditions,” 83-85.

• Sarah Allison, Ryan Heuser, Matthew Jockers, Franco Moretti, Michael Witmore. “Quantitative Formalism: An Experiment.” Stanford Literary Lab: Pamphlet 1. http://litlab.stanford.edu/?page_id=255

• Tanya Clement, “A thing not beginning and not ending?: Using Digital Tools to Distant-Read Gertrude Stein’s *The Making of Americans*,” *Literary and Linguistic Computing* 23 (2008): 361-81. (Moodle)

Praxis: Corpus comparisons using MONK <https://monk.library.illinois.edu/cic/public/> or Voyant voyant-tools.org More instructions will be forthcoming.

February 13th. Maps, networks, timelines.

• Franco Moretti: *Graphs, Maps, Trees*, pp. 35-64.

• Jo Guldi, “What is the Spatial Turn?,” *Spatial Humanities*, <http://spatial.scholarslab.org/spatial-turn/what-is-the-spatial-turn/>

• Matthew Wilkens, “Geolocation extraction and mapping of nineteenth-century American fiction,” DHCS 2011. <http://chicagocolloquium.org/dhcs-2011-program/>

• Croxall, Brian. “All Things Google: Google Maps.” *Profhacker*, April 5, 2011. <http://chronicle.com/blogs/profhacker/all-things-google-google-maps-labs/32421>.

• Wikipedia, “Social Network Analysis”; Elijah Meeks, “The Weird Geometry of the Internet,” <https://dhs.stanford.edu/social-media-literacy/tvtropes-pt-1-the-weird-geometry-of-the-internet/>

Praxis: Set up R to be ready for next week. I will provide instructions.

February 20th. Nuts and bolts.

Theory (sort of): • Stefan Gries, *Quantitative Corpus Linguistics with R*, “An Introduction to R,” pp. 19-68.

• Dan Cohen and Roy Rosenzweig, “To Mark Up, or Not to Mark Up.” In *Digital History: A Guide to Gathering, Preserving, and Presenting the Past on the Web*. University of Pennsylvania Press, 2005. <http://chnm.gmu.edu/digitalhistory/digitizing/3.php>

Praxis: Coding assignments, to include the classic “Hello, world,” and more!

February 27th. More nuts and bolts.

• *Quantitative Corpus Linguistics with R*, 68-103.

Praxis: How to produce your own visualizations using R.

March 5th. Getting the archive you need, in a form you can use.

Theory: • Cohen, Daniel J. “From Babel to Knowledge: Data Mining Large Digital Collections.” *D-Lib Magazine* 12, no. 3 (March 2006).

<http://www.dlib.org/dlib/march06/cohen/03cohen.html>.

• Ben Schmidt, “Technical notes,” *Sapping Attention*, 2011.

<http://sappingattention.blogspot.com/2011/02/technical-notes.html>

• *Quantitative Corpus Linguistics with R*, 7-17. (Note especially that “a word” is not a well defined concept!)

• Possible guest appearance by Jordan Sellers, who just finished building a massive collection that you’ll be using.

Praxis: Check out the Internet Archive, HathiTrust, and Gutenberg Archive. Which of these would be easiest to use? If the resources you need aren’t available in any of these places, where might you get them? Using either one of the corpora I have made available, or a set of documents you download from one of these sources, convert .txt files into a sparse table. Use a ruleset to normalize your data. Preserve metadata.

March 12th. Topic modeling.

Theory: Wikipedia, “Distributional Hypothesis.”

• Cameron Blevins, “Topic Modeling Martha Ballard’s Diary.”

<http://historying.org/2010/04/01/topic-modeling-martha-ballards-diary/>

- Robert Nelson, *Mining the Dispatch*. <http://dsl.richmond.edu/dispatch/>
- Ted Underwood, “18c tree,” plus a general comparison of Bayesian and non-Bayesian approaches to topic modeling that will probably be up on my blog in early Feb.
- David Blei and John D. Lafferty, “Topic Models,” 2009. (Moodle). Yes you can read it. Ignore the squiggles for now.

Praxis: Experiment with topic modeling tools I will provide.

March 19th. Spring break.

March 26th. Fly! Be free!

Theory: Wang, Xuerui, and McCallum, Andrew. “Topics Over Time: a non-Markov continuous-time model of topical trends.” *Proceedings of the Conference on Knowledge Discovery and Data Mining*, 2006. It’s on the web. Yes, you can read it. Ignore the squiggles.

Praxis: Using the tokenizer I’ve provided, transform a corpus of your choice into a sparse table. Run the sparse table through a topic modeling tool of your choice. Import the results into R, and visualize the topics as tree graphs. Also graph the frequency of each topic over time. Try creating a graph that superimposes the frequencies of two different topics. Or, if you prefer, do something different.

Critique of praxis: What did you discover?

April 2nd. Unit graduate conference on “Technology in Theory and Practice.”

I’ll probably cancel class, since it coincides with a Unit conference very much on our theme. Instead we’ll devote the week to individualized conferences, where I help you find the resources you need to explore a topic in your area of specialization and/or help you debug. Bugs are to be expected.

April 9th. Significance.

- xkcd, “Significance”
- Gries, *Quantitative Corpus Linguistics with R*, “Some Statistics for Corpus Linguistics,” pp. 173-217.
- Wikipedia, “Multiple comparisons”

Praxis: Using the topic models you generated on March 26th, look for topics that correlate across time. Is this correlation significant?

April 16th. Digital humanities as a social phenomenon.

From *Debates in the Digital Humanities*:

Tom Scheinfeldt, "Why Digital Humanities is 'Nice'," p. 59.

Bethany Nowviskie, "What Do Girls Dig," p. 235.

Alan Liu, "Where is Cultural Criticism in the Digital Humanities," p. 490.

Dave Parry, "The Digital Humanities or a Digital Humanism," p. 429.

Julia Flanders, "Time, Labor, and 'Alternate Careers'," p. 292

Amy Earheart, "Can Information Be Unfettered?," p. 309.

- Stephen Ramsay. "On Building," 2011. <http://lenz.unl.edu/papers/2011/01/11/on-building.html>

Praxis: Explore the Wordpress site. Suggest a reading for next week.

April 23rd. Blogging.

Theory: Readings you have suggested. In this field, it's likely there will be unexpected news by April we have to discuss.

Praxis: Write a blog post describing problems you are encountering or solving in your research.

April 30th. "Louis, I think this is the beginning of a beautiful friendship."

Theory: Blog posts written by everyone in the class. Possibly a case study of constructive opposition in the field, like the Moretti-Trumpener exchange in *Critical Inquiry* or the Wilkens-Rosen exchange in *Post45*. Probably the latter.

Praxis: Walk off into the fog to join the resistance (to whatever you like).

Resources I haven't been able to place on a specific day:

- The Blackwell Companion to Digital Humanities is a valuable resource. See especially Susan Hockey, "The History of Humanities Computing," http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-2-1&toc.depth=1&toc.id=ss1-2-1&brand=9781405103213_brand

- Lisa Spiro has several great lists of starting points for exploration:

<https://digitalresearchtools.pbworks.com/w/page/17801672/FrontPage>

<http://digitalscholarship.wordpress.com/2011/10/14/getting-started-in-the-digital-humanities/>