

## Cultural and Literary Text Mining

Kent K. Chang

(MSc student in Digital Humanities, DIS, UCL)

<kai-hsiung.chang.17@ucl.ac.uk>

August 9, 2018 · CALTMIG GLORIA (version 4.188)

### **COURSE OVERVIEW**

Welcome to Cultural and Literary Text Mining. As suggested by its title, this study group is where we learn and explore how to apply computational methods to the study of literature and culture. We look at the methodological debate of this “quantitative turn”, the tools and practices of which are becoming fundamental to the field now known as digital humanities. How does text mining inform literary and cultural studies? In a discipline that privileges close reading, what do people think about computational methods? What do we need to know to conduct data-driven research on literature and culture? What does a typical research project involve? We seek to find some answers to those questions here.

Although it’s called a study group, you may very well think of it as a student-initiated (Kent being the student), student-led, 15-credit postgraduate taught MA + MSc module. We will proceed in form of:

- **mini-lecture and seminar (theory)**, where a participant (Kent by default) tries to highlight key arguments in readings for each session, and other participants may share their opinions on and questions for them (if they have read them in advance, which is not strictly required)
- **Python workshop (practical)**, in which we first learn the basics of Python by making connections with JavaScript/PHP, and then try some tasks typically involved in a text mining project.

I imagine we will go through two major phases:

**Phase I** think about methodology and research questions + basic Python and statistics

**Phase II** look at some worked projects + practical experience of text mining and natural language processing in Python

**PREREQUISITE-ISH**

This group is open primarily to students currently doing an MA/MSc in Digital Humanities, as well as to those in the Department of Information Studies, at University College London. Typical participants have a BA degree in a humanities discipline such as English, history, classics, etc., and/or are interested in learning about or applying computational methods in the humanities.

The following modules are informal (or, indeed, nothing is formal for an inherently informal study group) prerequisites:

- INSTG008 Digital Resources in the Humanities
- G018 Introduction to Programming and Scripting

G008 has given us some idea of the scope of today's digital humanities research, and to a certain extent, this study group takes some of its sessions related text analysis as point of departure. G018, on the other hand, has equipped its students with programming concepts that are language independent, as well as some hands-on experience of coding—both are useful when you are starting to learn a new programming language (Python, in our case).

Having said that, you are likely to familiarize yourself with the nature of DH/humanities research and develop intuition for programming along the way.

**TIME AND PLACE**

Details will be announced during the first meeting; ideally we will meet at least once a week, in an evening during the week.

Unless otherwise noted, our meetings should take place in Foster Court G 3 1. Since this is not a computer cluster room, please bring your own laptops to the meetings, or borrow one at Science Library (UCL ID required).

**WEBSITE AND FACEBOOK GROUP**

Schedule and provisional agenda for each meeting will be made available at <https://caltmig.kentchang.com>. On this website you will also find links to lecture slides, reflective blog posts, Jupyter Notebooks, etc., where relevant. If you are still using Facebook, you may want to join our group at <https://www.facebook.com/groups/2079890148964442/>.

**TENTATIVE TOPICS**

The following is an indicative list of topics and readings. Readings listed below will be covered and discussed during mini-lectures and seminars; a secondary reading list will appear in Kent's reflective blog posts after (presumably) each session.

I don't expect you to finish all the readings—let's face it: even if you have time you don't necessarily want to spend them on reading, however dedicated you are to the subject matter; and admittedly, this looks like a lot of work. Ergo I am using the following editorial conventions to help you as you prepare for each session:

- REQUIRED : essential to follow the mini-lectures and seminars
- !! : highly recommended
- ! : recommended
- titles enclosed in squared brackets []: mentioned in the passing

And please understand that this list is subject to change.

## TOPIC 1 ■ Orientation

*What is CALTMIG?*

### Session 1

#### MODULE 1. Digital Scholarship in Literary and Cultural Studies

- CLUSTER A / GENERAL INTRODUCTION AND BACKGROUND
  - General introduction: Jockers and Underwood, “Text-Mining the Humanities”
  - Context for different modes of criticism:
    - \* Moretti, “Conjectures on World Literature”
    - \* Casanova, “Literature As a World”
    - \* Burrows, *Computation into Criticism*: Introduction
  - Towards a methodology:
    - \* !! Piper, “Think Small”
    - \* !! Sculley and Pasanek, “Meaning and Mining”
    - \* !! Underwood, “It Looks like You're Writing an Argument against Data in Literary Study . . .”
  - Case studies:
    - \* Kent Chang, “Articulating the ‘Love that Dare Not Speak Its Name’” (MS), Mapping Broadway [see also Miller, “Average Broadway”]
    - \* !! Long and So, “Literary Pattern Recognition”

---

*Phase I: Theoretical Foundation and Python Basics*


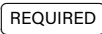
## TOPIC 2 ■ Scale and the “Quantitative Turn”:

### Rethinking the Part–Whole Dialectic in Literary Studies

*What is this quantitative approach to literary texts?*

## Session 2

### MODULE 2. *Scalar reading, pars prima*

- CLUSTER B / “CLOSE READING”, AN OVERVIEW
  - New Criticism:
    - \*  Eliot, “Tradition and the Individual Talent” (probably also his “Wasteland”)
    - \* Richards, *Practical Criticism*: introductory
    - \* Empson, *Seven Types of Ambiguity*: preface and chapter 1
  - Surface and Symptomatic Reading: Best and Marcus, “Surface Reading: An Introduction”
- CLUSTER C / RECONSIDERING SCALES: “THE CRISIS OF LARGENESS”
  -  English and Underwood, “Shifting Scales”
  - Williams, “The New Modesty in Literary Criticism”


#### Lab Session A

#### Python Fundamentals [o]: Installation and Setup

- Background reading: Saskar pp. 51–65
  - Getting to Know Python
  - Installation and Setup
  - Brief intro: request and NLTK

## Session 3

### MODULE 3. *Scalar reading, pars secunda*

- CLUSTER D / ALGORITHMS AND HUMANITIES RESEARCH
  - English and Underwood, “Shifting Scales” (reprise)
  - Underwood, “Theorizing Research Practices We Forgot to Theorize Twenty Years Ago”
  - Ramsay, *Reading Machines*: chap. 1
- CLUSTER E / SOME DEFINITIONAL ARTICULATIONS
  - Distant reading: Moretti, *Distant Reading*: “Slaughter House of Literature”, “Conjectures on World Literature”; *Graphs, Maps, Trees*: “Graphs, Maps, Trees”, “Graphs” [see also Walkowitz, *Born Translated*: chap. 1, and Goodwin and Holbo, *Reading Graphs, Maps, Trees*]
  - Cultural Analytics:  Piper, “There Will Be Numbers”
  - Macroanalysis: Jockers, *Macroanalysis*: part 1
  - [See also: Wilkens, “Canons, Close Reading, and the Evolution of Method”]

## Lab Session B

## Python Fundamentals [1]: Libraries, Data Type, and Controlling Workflow

- Optional homework: Think about what you want to study and mine.
- Background reading: Sarkar: chap. 1 (NLP Basics) and chap. 4 (Python)
- Dataquest: Python Basics
  - Files and Loops; Booleans and If Statements
  - List Operations; Dictionaries

## TOPIC 3 ■ From Tools to Methodologies

*How do we actually address humanities inquiries through computational methods (i.e. do something more than counting)?*

## Session 4-5

## MODULE 4. Think Searches, or How to Make DH Look Better

- CLUSTER F / SOME GENEALOGIES
  - REQUIRED Underwood, “A Genealogy of Distant Reading” [see also Underwood, *Why Literary Periods Mattered*: chap. 6]
  - ! Buurma and Heffernan, “Search and Replace”
  - Goldstone, “The *Doxa* of Reading” [see also Wimsatt Jr. and Beardsley, “The Intentional Fallacy”]
- CLUSTER G / SOME CRITICAL STANCES
  - Klein, “Distant Reading after Moretti”
  - Rhody, “Why I Dig”
  - Clement, “The Ground Truth of DH Text Mining”
  - Golumbia, “Death of a Discipline” [see also Spivak, *Death of a Discipline*: chap. 1, and Derrida, *Limited Inc.*: e.g. “Signature Event Context”]
  - Liu, “Where Is Cultural Criticism in the Digital Humanities?”
  - [Marche](#), “Literature Is Not Data”

## Lab Session C

**Python Fundamentals [2]:** Revision; Functional and Modular Programming

- Background reading:
  - Sarkar: chap. 1 (NLP Basics), chap. 4, pp. 84–91 (Functional Programming), and chap. 4, pp. 91–105 (to the end of chapter)
  - Guttag, *Introduction to Computation and Programming Using Python*: chap. 4

---

- Self-directed study: Dataquest: Python Basics
  - Functions


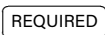
## Lab Session D

**Python Fundamentals [3]:** Obtaining Texts

- Optional homework: Prepare to share your project ideas with us.
- Downloading texts from Project Gutenberg using Corpus-DB
- Other sources: e.g. HathiTrust and .txtLAB@McGill
- Web Scraping and BeautifulSoup

---

**Session 6**

- CLUSTER H / DH IN POPULAR MEDIA
  -  Nyhan, “It Is Time to Address the Public Communication of DH”
  - Piper, “Why Are Non-Data Driven Representations of Data-Driven Research in the Humanities So Bad?”
  - Some case studies:
    - \*  academics writing for non-academics:
      - Piper and So, “Women Write About Family, Men Write About War” (*The New Republic*) [see also: Underwood, “The Real Problem with Distant Reading”]
      - Sinykin, “How Computational Analysis Is Teaching Us to Read in New Ways” (*The Washington Post*)
    - \* media coverage of Underwood’s research:
      - Flood, “Women Better Represented in Victorian Novels than Modern, Finds Study” (*The Guardian*)
      - Eschner, “Women Were Better Represented in Victorian Novels Than Modern Ones” (*Smithsonian*)

- \* response to Jockers’s *The Bestseller Code*
  - Tolentino, ““[The Bestseller Code](#)” Tells Us What We Already Know” (*The New Yorker*)
  - Althoff, “[Algorithms Could Save Book Publishing—but Ruin Novels](#)” (*Wired*)
- \* [See also: the Pinker–Wieseltier debate: [Pinker, “Science Is Not Your Enemy”](#), [Wieseltier, “Crimes Against Humanities”](#), and [Pinker and Wieseltier, “Science vs. the Humanities, Round III”](#)]

#### MODULE 5. Think Future, or How to Teach DH

- CLUSTER I / CURRICULA AND PEDAGOGIES
  - REQUIRED Underwood, “A Broader Purpose”
  - !! Goldstone, “Teaching Quantitative Methods”
  - Ramsay, “Digital Humanities Pedagogy”
  - So (@richardjeanso), [Twitter Post](#) from August 2, 2018

#### Lab Session E

##### Statistics Primer with Python [o]

- numpy, vectors, and matrices
- Essentials of probability and Bayes’ Rule

#### Sessions 7-8

#### MODULE 6. “Think Small”, or How to Deal with Numbers

- CLUSTER J / LITERARY DATA
  - [Marche, “Literature Is Not Data”](#) (reprise)
  - Gitelman, “*Raw Data*” Is an Oxymoron: introduction
  - Bode, “The Equivalence of “Close” and “Distant” Reading”
- CLUSTER K / LITERARY MEASUREMENTS
  - Moretti, ““Operationalizing””
  - Hoover, “Quantitative Analysis and Literary Studies”
  - Miller, “Average Broadway” (reprise)
- CLUSTER L / LITERARY MODELING
  - Geoffrey Rockwell and Stefan Sinclair, “A Model Theory”
  - McCarty, “Knowing . . . : Modeling in Literary Studies”
  - REQUIRED Piper, “Think Small” (reprise)
  - !! So, ““All Models Are Wrong””
  - Drucker, “Why Distant Reading Isn’t”

## Lab Session F

## Statistics Primer with Python [1–2]

- panda and matplotlib
- Bayes' Rule cont'd (prior and posterior)
- discrete and continuous random variables

---

*Phase II: Cultural/Literary Text Mining in Practice*

## TOPIC 4 ■ Text Classification

## Session 9

## MODULE 7. Stylometrics, Some Worked Projects

- CLUSTER M / AUTHORSHIP ATTRIBUTION AND STYLE
  - Spitzer, *Linguistics and Literary History*: chap. 1
  - Burrows, *Computation into Criticism*: Introduction (reprise)
  - Jockers and Underwood, “Text-Mining the Humanities” (reprise)
  - Jockers, *Macroanalysis*: chap. 6 (Style)
- CLUSTER N / GENDER AND STYLE
  - Clement, “The Ground Truth of DH Text Mining” (reprise)
  - Butler, *Gender Trouble*: at least chap. 1 (Subjects of Sex/Gender/Desire)
  - Argamon et al., “Gender, Genre, and Writing Style in Formal Written Texts”

---

Sessions 10–11

## MODULE 8. Cultural Studies, Some Worked Projects

- CLUSTER O / CULTURAL CAPITAL AND LITERARY FIELD
  - REQUIRED Bourdieu, *The Field of Cultural Production*: ideally pt. 1, at least chap. 1
  - !! Miller, “Average Broadway” (reprise 2)
  - !! [Piper and Portelance, “How Cultural Capital Works”](#)
  - Algee-Hewitt et al., “Canon/Archive: Large-Scale Dynamics in the Literary Field”
  - [See also: Quist, “Laurelled Lives”]




**Lab Session G****Text Mining and NLP with Python [0]:****Text Normalization and Feature Engineering**

- NLP Basics—Sarkar: chap. 1 and 3 (Natural Language Basics; Processing and Understanding Text)

**Lab Session H****Text Mining and NLP with Python [1]:****Text Classification**

- Text Classification—Sarkar: chap. 4 (Text Classification)

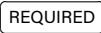
**TOPIC 5 ■ Social Network Analysis****Session 12****MODULE 9. Social Network Analysis**

- CLUSTER P / SOCIAL NETWORK ANALYSIS
  - Background: [Weingart, “Demystifying Networks”](#)
  -  So and Long, “Network Analysis and the Sociology of Modernism”
  - Schich et al., “A Network Framework of Cultural History”


**Lab Session I****Text Mining and NLP with Python [2]: Network Analysis with **NetworkX****

- Al-Taie (not Sarkar!): chap. 2

**TOPIC 6 ■ Topic Modeling****Session 13****MODULE 10. Topic modeling**

- CLUSTER Q / TOPIC MODELING: BACKGROUND
  -  [Underwood, “Topic Modeling Made Just Simple Enough.”](#)
  - Blei, “Probabilistic Topic Models” (alert: a bit mathy)

**Session 14**

- CLUSTER R / TOPIC MODELING: APPLICATIONS IN THE HUMANITIES
  -  Jockers, *Macroanalysis*: chap. 8
  - Jockers and Mimno, “Significant Themes in 19th-Century Literature”

- !! Goodwin, “Topic Modeling Modernism/Modernity”
- !! Erlin, “Topic Modeling, Epistemology, and the English and German Novel”
- Goldstone and Underwood, “The Quiet Transformations of Literary Studies” [see also: Foucault, “The Discourse on Language”]

#### Lab Session J

#### Text Mining and NLP with Python [3]: Topic Modeling

- Sarkar: chap. 5

## TOPIC 7 ■ Vector Semantics

### Session 15

#### MODULE 11. Vector Semantics

- CLUSTER S / VECTOR SEMANTICS
  - REQUIRED Gavin, “Vector Semantics, William Empson, and the Study of Ambiguity”
  - Turney and Pantel, “From Frequency to Meaning: Vector Space Models of Semantics”
  - Bamman, Underwood, and Smith, “A Bayesian Mixed Effects Model of Literary Character”

#### Lab Session K

#### Text Mining and NLP with Python [4]: word2vec

- [Historical Word2Vec](#)

## TOPIC 8 ■ Commencement

*Let's start from here.*

### Session 16

#### MODULE 12. More methods

- CLUSTER T / MORE METHODS
  - Sentiment analysis—Stanford Literary Lab, “Mapping London’s Emotions”; *Optional practice*: see Sarkar chap. 7
  - Geospatial analysis—Wilkens, “The Geographic Imagination of Civil War-Era American Fiction”
  - Mixed methods—Himmelboim, McCreery, and Smith, “Birds of a Feather Tweet Together”

#### MODULE 13. Finale: concluding remarks

## TENTATIVE PRIMARY READING LIST

- Algee-Hewitt, Mark, Sarah Allison, Marissa Gemma, Ryan Heuser, Franco Moretti, and Hannah Walser. "Canon/Archive: Large-Scale Dynamics in the Literary Field." January 2016. <https://litlab.stanford.edu/LiteraryLabPamphlet11.pdf>.
- Argamon, Shlomo, Moshe Koppel, Jonathan Fine, and Anat Rachel Shimoni. "Gender, Genre, and Writing Style in Formal Written Texts." *TEXT* 23 (2003): 321–46.
- Bamman, David, Ted Underwood, and Noah A. Smith. "A Bayesian Mixed Effects Model of Literary Character," 370–79. Association for Computational Linguistics, 2014.
- Best, Stephen, and Sharon Marcus. "Surface Reading: An Introduction." *Representations* 108, no. 1 (23 2009): 1–21.
- Blei, David M. "Probabilistic Topic Models." *Communications of the ACM* 55, no. 4 (2012): 77–84.
- Bode, Katherine. "The Equivalence of "Close" and "Distant" Reading; or, Toward a New Object for Data-Rich Literary History." *Modern Language Quarterly* 78, no. 1 (March 2017): 77–106.
- Bourdieu, Pierre. *The Field of Cultural Production: Essays on Art and Literature*. Reprinted. Edited by Randal Johnson. Cambridge: Polity Press, 1993.
- Burrows, J. F. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Oxford University Press, 1987.
- Butler, Judith. *Gender Trouble: Feminism and the Subversion of Identity*. New York: Routledge, 1990.
- Buurma, Rachel Sagner, and Laura Heffernan. "Search and Replace: Josephine Miles and the Origins of Distant Reading." April 11, 2018. <https://modernismmodernity.org/forums/posts/search-and-replace>.
- Casanova, Pascale. "Literature As a World." *New Left Review* 31 (2005): 71–90.
- Clement, Tanya E. "The Ground Truth of DH Text Mining." In *Debates in the Digital Humanities 2016*, redacted by Matthew K. Gold and Lauren F. Klein, 534–35. University of Minnesota Press, 2016.
- Derrida, Jacques. *Limited Inc*. Evanston, IL: Northwestern University Press, 1988.
- Drucker, Johanna. "Why Distant Reading Isn't." *PMLA* 132, no. 3 (May 2017): 628–35.
- Eliot, T. S. "Tradition and the Individual Talent." In *The Norton Anthology of Theory and Criticism*, 1st ed, edited by Vincent B. Leitch, 1092–98. New York: Norton, 2001.
- Empson, William. *Seven Types of Ambiguity*. [3. ed.], 14. print. New Directions paperback 204. New York: New Directions Publ, 1984.
- English, James F., and Ted Underwood. "Shifting Scales: Between Literature and Social Science." *Modern Language Quarterly* 77, no. 3 (September 2016): 277–95.
- Erlin, Matt. "Topic Modeling, Epistemology, and the English and German Novel." *Journal of Cultural Analytics*, May 1, 2017. Accessed July 3, 2018.
- Foucault, Michel. "The Discourse on Language." In *Truth*, edited by Jos Medina and David Wood, 315–35. Ames, Iowa, USA: Blackwell Publishing, January 1, 2005.
- Gavin, Michael. "Vector Semantics, William Empson, and the Study of Ambiguity." *Critical Inquiry* 44, no. 4 (June 2018): 641–73.
- Geoffrey Rockwell and Stefan Sinclair. "A Model Theory: Thinking Through Hermeneutical Things." In *Hermeneutica: Computer-Assisted Interpretation in the Humanities*, 149–66. MIT Press, 2016.

- Gitelman, Lisa, ed. *“Raw Data” Is an Oxymoron*. Infrastructures series. Cambridge, Massachusetts: The MIT Press, 2013.
- Goldstone, Andrew. “Teaching Quantitative Methods: What Makes It Hard (in Literary Studies).” In *Debates in the Digital Humanities 2018*, edited by Matthew K. Gold and Lauren F. Klein. Minneapolis: University of Minnesota Press, 2018.
- . “The *Doxa* of Reading.” *PMLA* 132, no. 3 (May 2017): 636–42.
- Goldstone, Andrew, and Ted Underwood. “The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us.” *New Literary History* 4, no. 2 (2014): 359–84.
- Golumbia, D. “Death of a Discipline.” *differences* 25, no. 1 (2014): 156–76.
- Goodwin, Johnathan. “Topic Modeling Modernism/Modernity.” May 26, 2016. <https://modernismmodernity.org/forums/posts/topic-modeling-modernismmodernity>.
- Goodwin, Jonathan, and John Holbo, eds. *Reading Graphs, Maps, Trees: Responses to Franco Moretti*. Anderson, SC: Parlor Press, 2011.
- Guttag, John. *Introduction to Computation and Programming Using Python*. Revised and expanded edition. Cambridge, Massachusetts: The MIT Press, 2013.
- Himmelboim, Itai, Stephen McCreery, and Marc Smith. “Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter.” *Journal of Computer-Mediated Communication* 18, no. 2 (January 2013): 40–60.
- Hoover, David L. “Quantitative Analysis and Literary Studies.” In *A Companion to Digital Literary Studies*, edited by Raymond George Siemens and Susan Schreibman, 517–53. Blackwell companions to literature and culture 50. Malden, MA: Blackwell Pub, 2007.
- Jockers, Matthew L. *Macroanalysis: Digital Methods and Literary History*. Urbana, IL: University of Illinois Press, 2013.
- Jockers, Matthew L., and David Mimno. “Significant Themes in 19th-Century Literature.” [pre-print], August 2012.
- Jockers, Matthew L., and Ted Underwood. “Text-Mining the Humanities.” In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 291–306. Chichester, UK: John Wiley & Sons, Ltd, November 27, 2015.
- Klein, Lauren. “Distant Reading after Moretti.” January 10, 2018. <http://klein.com/2018/01/distant-reading-after-moretti/>.
- Liu, Alan. “Where Is Cultural Criticism in the Digital Humanities?” In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 490–510. University of Minnesota Press, January 1, 2012.
- Long, Hoyt, and Richard Jean So. “Literary Pattern Recognition: Modernism between Close Reading and Machine Learning.” *Critical Inquiry* 42, no. 2 (January 2016): 235–67.
- Marche, Stephen. “Literature Is Not Data: Against Digital Humanities.” October 28, 2008. <https://lareviewofbooks.org/article/literature-is-not-data-against-digital-humanities/>.
- McCarty, Willard. “Knowing . . . : Modeling in Literary Studies.” In *A New Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth, 391–401. Chichester, UK: John Wiley & Sons, Ltd, November 27, 2015.
- Miller, Derek. “Average Broadway.” *Theatre Journal* 68, no. 4 (2016): 529–53.
- Moretti, Franco. “Conjectures on World Literature.” *New Left Review* 1 (2000): 54–68.
- . *Distant Reading*. London ; New York: Verso, 2013.
- . *Graphs, Maps, Trees: Abstract Models for a Literary Theory*. London: Verso, 2005.

- Moretti, Franco. "Operationalizing": Or, the Function of Measurement in Modern Literary Theory." December 2013. <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf>.
- Nyhan, Julianne. "It Is Time to Address the Public Communication of DH." *Digital Humanities Quarterly* 10, no. 3 (2016).
- Pinker, Stephen. "Science Is Not Your Enemy." August 7, 2013. <https://newrepublic.com/article/114127/science-not-enemy-humanities>.
- Pinker, Steven, and Leon Wieseltier. "Science vs. the Humanities, Round III." September 27, 2013. <https://newrepublic.com/article/114754/steven-pinker-leon-wieselstier-debate-science-vs-humanities>.
- Piper, Andrew. "There Will Be Numbers." *Cultural Analytics*, May 23, 2016.
- . "Think Small: On Literary Modeling." *PMLA* 132, no. 3 (May 2017): 651–58.
- . "Why Are Non-Data Driven Representations of Data-Driven Research in the Humanities So Bad?" September 17, 2017. <https://txtlab.org/2017/09/why-are-non-data-driven-representations-of-data-driven-research-in-the-humanities-so-bad/>.
- Piper, Andrew, and Eva Portelance. "How Cultural Capital Works: Prizewinning Novels, Bestsellers, and the Time of Reading." May 10, 2016. <http://post45.research.yale.edu/2016/05/how-cultural-capital-works-prizewinning-novels-bestsellers-and-the-time-of-reading/>.
- Piper, Andrew, and Richard Jean So. "Women Write About Family, Men Write About War." April 8, 2016. <https://newrepublic.com/article/132531/women-write-family-men-write-war>.
- Quist, Jennifer. "Laurelled Lives: The Swedish Academy's Praise for Its Prizewinners." *New Left Review* 104 (March–April 2017): 93–106.
- Ramsay, Stephen. "Programming with Humanists: Reflections on Raising an Army of Hacker-Scholars in the Digital Humanities." In *Programming with Humanists*, edited by Brett D Hirsch, 217–40. OCLC: 1043073947. Cambridge: Open Book Publishers, 2012.
- . *Reading Machines: Toward an Algorithmic Criticism*. University of Illinois Press, 2011. Accessed April 23, 2018.
- Rhody, Lisa Marie. "Why I Dig: Feminist Approaches to Text Analysis." In *Debates in the Digital Humanities 2016*, redacted by Matthew K. Gold and Lauren F. Klein, 536–39. University of Minnesota Press, 2016.
- Richards, I. A. *Practical Criticism: A Study of Literary Judgment*. Place of publication not identified: publisher not identified, 1929.
- Schich, M., C. Song, Y.-Y. Ahn, A. Mirsky, M. Martino, A.-L. Barabasi, and D. Helbing. "A Network Framework of Cultural History." *Science* 345, no. 6196 (August 1, 2014): 558–62.
- Sculley, D., and B. M. Pasanek. "Meaning and Mining: The Impact of Implicit Assumptions in Data Mining for the Humanities." *Literary and Linguistic Computing* 23, no. 4 (September 29, 2008): 409–24.
- Sinykin, Dan. "How Computational Analysis Is Teaching Us to Read in New Ways." July 30, 2018. [https://www.washingtonpost.com/news/posteverything/wp/2018/07/30/how-computational-analysis-is-teaching-us-to-read-in-new-ways/?utm\\_term=.6bd5594ec88c](https://www.washingtonpost.com/news/posteverything/wp/2018/07/30/how-computational-analysis-is-teaching-us-to-read-in-new-ways/?utm_term=.6bd5594ec88c).
- So, Richard Jean. "'All Models Are Wrong.'" *PMLA* 132, no. 3 (May 2017): 668–73.
- . "Twitter Post." August 2, 2018, 6:36am (BST). <https://twitter.com/RichardJeanSo/>.
- So, Richard Jean, and Hoyt Long. "Network Analysis and the Sociology of Modernism." *boundary 2* 40, no. 2 (June 1, 2013): 147–82.
- Spitzer, Leo. *Linguistics and Literary History*. NJ: Princeton University Press, 1948.

- Spivak, Gayatri Chakravorty. *Death of a Discipline*. The Wellek Library lectures in critical theory. New York Chichester, West Sussex: Columbia University Press, 2003.
- Stanford Literary Lab. "Mapping London's Emotions." *New Left Review* 101 (September–October 2016): 63–91.
- Al-Taie, Mohammed Zuhair. *Python for Graph and Network Analysis*. New York, NY: Springer Berlin Heidelberg, 2017.
- Turney, Peter D., and Patrick Pantel. "From Frequency to Meaning: Vector Space Models of Semantics." *CoRR* abs/1003.1141 (2010).
- Underwood, Ted. "A Broader Purpose." January 4, 2018. <https://tedunderwood.com/2018/01/04/a-broader-purpose/>.
- . "A Genealogy of Distant Reading." *Digital Humanities Quarterly* 11, no. 2 (2017).
- . "It Looks like You're Writing an Argument against Data in Literary Study . . ." September 21, 2017. <https://tedunderwood.com/2017/09/21/it-looks-like-youre-writing-an-argument-against-data-in-literary-study/>.
- . "The Real Problem with Distant Reading." May 29, 2016. <https://tedunderwood.com/2016/05/29/the-real-problem-with-distant-reading/>.
- . "Theorizing Research Practices We Forgot to Theorize Twenty Years Ago." *Representations* 127, no. 1 (August 2014): 64–72.
- . "Topic Modeling Made Just Simple Enough." April 7, 2012. <https://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/>.
- . *Why Literary Periods Mattered: Historical Contrast and the Prestige of English Studies*. California: Stanford UP, 2013.
- Walkowitz, Rebecca L. *Born Translated: The Contemporary Novel in an Age of World Literature*. Literature now. New York: Columbia University Press, 2015.
- Weingart, Scott B. "Demystifying Networks." December 14, 2011. <http://scottbot.net/lets-talk-about-networks/>.
- Wieseltier, Leon. "Crimes Against Humanities." September 4, 2013. <https://newrepublic.com/article/114548/leon-wieseltier-responds-steven-pinkers-scientism>.
- Wilkens, M. "The Geographic Imagination of Civil War-Era American Fiction." *American Literary History* 25, no. 4 (December 1, 2013): 803–40.
- Wilkens, Matthew. "Canons, Close Reading, and the Evolution of Method." In *Debates in the Digital Humanities*, edited by Matthew K. Gold, 249–58. University of Minnesota Press, January 1, 2012.
- Williams, Jeffrey J. "The New Modesty in Literary Criticism." January 5, 2015. <https://www.chronicle.com/article/The-New-Modesty-in-Literary/150993>.
- Wimsatt Jr., W. K., and M. C. Beardsley. "The Intentional Fallacy." *The Sewanee Review* 54, no. 3 (Jul.-Sep., 1946): 468–88.

## ACKNOWLEDGEMENTS

This document is created based on Andrew Goldstone's  $\LaTeX$  template at <https://github.com/agoldst/tex/blob/master/syllabus/example/syllabus.tex>.

In preparing for this document, I have referred extensively to syllabi publicly available online. Those made by Andrew Goldstone (350:509, 350:596), Andrew Piper (LLCU 614, LLCU 255), David Mimno (INFO 3350), Mark Algee-Hewitt (ENGLISH 184E, ENGLISH 253, ENGLISH 354), and Matthew Jockers (ENGLISH 4/898) have significantly informed my compilation.