

# English 184E | Literary Text Mining

Autumn, 2019-2020

## Details

Instructor: Erik Fredner  
Email: fredner@stanford.edu  
Time: M & W, 10:30-11:50  
Place: Encina West 202  
Office: Margaret Jacks Hall (Building 460), Room 419  
Office Hours: Fridays, 1-3pm

## Description

Computational methods have made it possible to analyze literature in new ways and at new scales. This course trains students in theories and methods of computational literary studies. It requires no background in computer programming or literary criticism. We begin with fundamentals of the Python programming language before moving on to computational analyses of literary texts. Our analyses will be informed by critical readings. While programming distinguishes this course from many other humanities classes, this course is fundamentally about literature, and what we can learn about it at the new scales afforded by computation.

## Objectives

Students will be able to:

- Plan and justify a program of computational literary analysis.
- Create a corpus and metadata to study a research question.
- Use all basic elements of Python proficiently.
- Use several data science and text mining packages.
- Visualize data.
- Integrate data analysis into literary criticism.

## Technology

All students will need a laptop running macOS, Linux, or Windows that they can bring to class. (MacOS is preferred.) Chromebooks and iPads will not work for this course. If you need access to an appropriate laptop, please let me know ASAP.

## Readings

All readings for this course will be posted on Canvas. Please note that many readings will be excerpted. If you need paper copies, please let me know ASAP.

## Calendar

I give abbreviated titles for some readings in the calendar. Full citations are available in the bibliography at the end of the syllabus.

| Date   | Reading  |
|--------|--|
| Sep 23 | Kirschenbaum, "What is 'digital humanities'?"          |
| Sep 25 | Liu, "The Meaning of the Digital Humanities"           |
| Sep 30 | Algee-Hewitt et al., "Between Canon and Corpus"        |
|        | Porter, "Popularity and Prestige"                      |
| Oct 2  | Burrows, <i>Computation into Criticism</i>             |
| Oct 7  | Piper, <i>Enumerations</i>                             |
| Oct 9  | Arnold and Tilton, "Role of Statistics in DH"          |
| Oct 14 | <i>No class</i>  |
| Oct 16 | Ramsay, "An Algorithmic Criticism"                     |
| Oct 21 | Bode, "Equivalence of 'Close' and 'Distant' "          |
| Oct 23 | Guldi, "Critical Search"                               |
| Oct 28 | Mandell, "Gender and Cultural Analytics"               |
|        | Underwood et al, "The Transformation of Gender"        |
| Oct 30 | Heuser and Le-Khac, "A Quantitative Literary History"  |
| Nov 4  | Fields, <i>Racecraft</i>                               |
|        | McPherson, "Why are the digital humanities so white?"  |
| Nov 6  | Ruberg et al., "Toward a Queer Digital Humanities"     |
| Nov 11 | Risam, "The Stakes of Postcolonial Digital Humanities" |
| Nov 13 | English, "Now, Not Now"                                |
|        | Blei, "Probabilistic Topic Models"                     |
| Nov 18 | Underwood, <i>Distant Horizons</i>                     |
| Nov 20 | D'Ignazio and Klein, "Feminist Data Visualization"     |
|        | Clement, "Text Mining, Visualization"                  |
|        | Drucker, "Graphical Display"                           |
| Nov 25 | <i>Recess</i>  |
| Nov 27 | <i>Recess</i>  |
| Dec 2  | Presentations  |
| Dec 4  | Presentations  |
| Dec 13 | <i>Final paper due</i>                                 |

## Literary Lab meetings

Students are encouraged to attend meetings of the Stanford Literary Lab, a research group that focuses on computational text analysis: <https://litlab.stanford.edu/events/>

## Grading

I round grades to the nearest integer.

| Letter | Range      |
|--------|------------|
| A      | 93 or more |
| A-     | 90-92      |
| B+     | 87-89      |
| B      | 83-86      |
| B-     | 80-82      |
| C+     | 77-79      |
| C      | 73-76      |
| C-     | 70-72      |
| D+     | 67-69      |
| D      | 60-66      |
| F      | 59 or less |

## Evaluation

| Item          | Percentage | Description  |
|---------------|------------|--|
| Homework      | 40%        | Weekly assignments that apply material from class.                                 |
| Participation | 20%        | Attendance, discussion, and lab work.  |
| Presentation  | 10%        | One in-class presentation of original research.                                    |
| Final Essay   | 30%        | Uses evidence from computational analysis to advance a literary critical argument. |

Homework will include programming practice, new problems, and written reflections on readings or results.

Attendance is crucial in this course since computational methods build on each other. If you must miss a class, I expect you to cover the material you missed during office hours, or with classmates.

Further details on the presentation and final essay will be circulated later in the quarter.

## **Miscellaneous Items**

### **Accommodations**

- Please let me know about any accommodations you require ASAP.
- Contact the Office of Accessible Education if you need an updated letter.

### **Absences**

- Students are allowed one unexcused absence.
- Email me at least 24 hours prior to class to secure an excused absence.
- If you know you will be absent on specific days this quarter (e.g. away games), please let me know as far in advance as possible.

### **Email**

- I usually reply to email within a day. If you have not heard from me after that time, please write again.

### **Office hours**

- If you cannot make my regular office hours, I would be happy to schedule a meeting at another time that works for both of us.
- To discuss writing in progress, I ask that you send me your text two days before we meet.

### **Late work**

- I apply a -10% penalty for each day that passes after the original deadline to late work.
- I do not accept work that is more than four days late.
- (Yes, weekends count.)

### **Other**

- Take advantage of the Hume Center for Writing and Speaking while drafting and revising your final essay.
- If you get stuck on a programming problem, I encourage you to reach out to your classmates to solve it together.
- Mental and physical health are essential for you to do your best work. If you are feeling unwell at any point during the quarter, please talk to me.

## **Appendix**

### **Honor Code**

The Honor Code is an undertaking of the students, individually and collectively:

- a. that they will not give or receive aid in examinations; that they will not give or receive unpermitted aid in class work, in the preparation of reports, or in any other work that is to be used by the instructor as the basis of grading;
- b. that they will do their share and take an active part in seeing to it that others as well as themselves uphold the spirit and letter of the Honor Code.

The faculty on its part manifests its confidence in the honor of its students by refraining from proctoring examinations and from taking unusual and unreasonable precautions to prevent the forms of dishonesty mentioned above. The faculty will also avoid, as far as practicable, academic procedures that create temptations to violate the Honor Code.

While the faculty alone has the right and obligation to set academic requirements, the students and faculty will work together to establish optimal conditions for honorable academic work.

### **Fundamental Standard**

Students are expected to show both within and without the University such respect for order, morality, personal honor, and the rights of others as is demanded of good citizens. Failure to do this will be sufficient cause for removal from the University.

### **Students with Documented Disabilities**

Students who may need an academic accommodation based on the impact of a disability must initiate the request with the Office of Accessible Education (OAE). Professional staff will evaluate the request with required documentation, recommend reasonable accommodations, and prepare an Accommodation Letter for faculty dated in the current quarter in which the request is being made. Students should contact the OAE as soon as possible since timely notice is needed to coordinate accommodations.

The Office of Accessible Education  
563 Salvatierra Walk  
650-723-1066  
<https://studentaffairs.stanford.edu/oae>

### **Campus Resources**

Hume Center for Speaking and Writing  
<https://sites.stanford.edu/undergrad/tutoring-support/hume-center>

Office of Sexual Assault & Relationship Abuse Education & Response (SARA)  
<https://sara.stanford.edu>

Counseling and Psychological Services (CAPS)

<https://vaden.stanford.edu/caps>

The Bridge Peer Counseling

<https://haas.stanford.edu/students/cardinal-commitment/bridge>

English for Foreign Students

<https://language.stanford.edu/programs/efs/languages/english-foreign-students>

Academic Skills Coaching

<http://learningconnection.stanford.edu/academic-skills-coaching>

Undergraduate Advising

<https://undergrad.stanford.edu/advising>

Community Center Resources

<https://undergrad.stanford.edu/tutoring-support/community-center-resources>

### **Acknowledgments**

This syllabus extends past iterations of this course taught by Mark Algee-Hewitt and Ryan Heuser. It has also benefitted, directly and indirectly, from syllabi by David Bamman, Andrew Goldstone, Laura McGrath, and J.D. Porter.

## Bibliography

Algee-Hewitt, Mark, and Mark McGurl. *Between Canon and Corpus: Six Perspectives on 20th-Century Novels*.

Arnold, Taylor, and Lauren Tilton. *New Data? The Role of Statistics in DH*. <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9acie3a60/section/a2a6a192-fo4a-4082-afaa-97c76a75b21c>. Accessed 21 Aug. 2019.

Blei, David M. 'Probabilistic Topic Models'. *Communications of the ACM*, vol. 55, no. 4, Apr. 2012, p. 77, doi:10.1145/2133806.2133826. Accessed 21 Aug. 2019.

Bode, Katherine. 'The Equivalence of "Close" and "Distant" Reading; or, Toward a New Object for Data-Rich Literary History'. *Modern Language Quarterly*, vol. 78, no. 1, 2017, pp. 77–106.

Burrows, John Frederick. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Clarendon Press, 1987.

D'Ignazio, Catherine, and Lauren F. Klein. *Feminist Data Visualization*. p. 5.

Drucker, Johanna. 'Humanities Approaches to Graphical Display'. *Digital Humanities Quarterly*, vol. 5, no. 1, 2011, pp. 1–21.

English, James F. 'Now, Not Now: Counting Time in Contemporary Fiction Studies'. *Modern Language Quarterly*, vol. 77, no. 3, 2016, pp. 395–418.

Fields, Karen E., and Barbara Jeanne Fields. *Racecraft: The Soul of Inequality in American Life*. Verso Books, 2014.

Gold, Matthew K. *Debates in the Digital Humanities*. U of Minnesota Press, 2012.

Guldi, Jo. 'Critical Search: A Procedure for Guided Reading in Large-Scale Textual Corpora'. *Journal of Cultural Analytics*, 2018.

Heuser, Ryan, and Long Le-Khac. 'A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method'. *Literary Lab Pamphlets*.

Liu, Alan. 'The Meaning of the Digital Humanities'. *Pmla*, vol. 128, no. 2, 2013, pp. 409–423.

Mandell, Laura. 'Gender and Cultural Analytics: Finding or Making Stereotypes?' *Debates in the Digital Humanities*, <https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9acie3a60/section/5d9c1b63-7b60-42dd-8cda-bde837f638f4>. Accessed 21 Aug. 2019.

McPherson, Tara. 'Why Are the Digital Humanities so White? Or Thinking the Histories of Race and Computation'. *Debates in the Digital Humanities*, <https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfbdie/section/20df8acd-9ab9-4f35-8a5d-e91a25f4a0ea>. Accessed 21 Aug. 2019.

Piper, Andrew. *Enumerations: Data and Literary Study*. University of Chicago Press, 2018.

Porter, J.D. 'Popularity/Prestige'. *Literary Lab Pamphlets*.

Ramsay, Stephen. *Reading Machines: Toward and Algorithmic Criticism*. University of Illinois Press, 2011.

Risam, Roopika. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Northwestern University Press, 2018.

Ruberg, Bonnie, et al. 'Toward a Queer Digital Humanities'. *Debates in the Digital Humanities*, <https://dhdebates.gc.cuny.edu/read/untitled-4e08b137-aec5-49a4-83c0-38258425f145/section/e1a72933-2aef-4158-b21c-e6fe35afco91>. Accessed 21 Aug. 2019.

Tenen, Dennis Yi. 'Toward a Computational Archaeology of Fictional Space'. *New Literary History*, vol. 49, no. 1, 2018, pp. 119–147.

Underwood, Ted. *Distant Horizons: Digital Evidence and Literary Change*. University of Chicago Press, 2019.

Underwood, Ted, et al. *The Transformation of Gender in English-Language Fiction*. 2018.