**The Text in the Machine: Digital Approaches to Literature**  
**English 153H**  
**Monday/Wednesday 3:30-5:20 PM, 240-110**

**Course Description**

Whether you’re an avid programmer and an equally avid science fiction reader or an English major who designs apps in her spare time, you’ve probably had some experience switching between literary and technological modes of thinking, between following a story and debugging a script. What if these distinct cognitive styles could be combined? How can our understanding of narrative be enriched through computational analysis — and how might applying empirical techniques to interpretive questions change the way we think about both realms?

In this course, we will try to learn things about texts by counting, clustering, and networking their contents, exploring the burgeoning field of digital humanities. We will practice some basic techniques of quantitative textual analysis, from identifying a text’s most distinctive words to determining its favorite topics and themes. But we will also find that these technical questions point us back to more basic interpretive or even philosophical ones: what is an author? What is a character? When we have a sense of a poem’s mood, what exactly are we responding to? When we feel suspense or anxiety while reading a story, how is the text creating that effect?

We will spend the first half of the quarter thinking about what it means to treat literature empirically, drawing upon recent research in digital humanities — including the work of Stanford’s own Literary Lab. The second half of the quarter will be dedicated to group projects in which you will ask and answer your own empirical questions about a large corpus of texts. By the end of the quarter, you and your partner(s) will have generated a hypothesis, tested it using quantitative methods, and analyzed the results in a short paper and presentation to the class.

**Course Goals**

In addition to introducing you to the field of digital humanities, this course is intended to help you practice working collaboratively on a humanities project, since almost all DH investigations involve some form of teamwork. By the end of the quarter, you should be able to:

1) Identify and explain major concepts in quantitative textual analysis (e.g. word frequency, word collocation, topic modeling, character networks), and execute at least one of these techniques independently.  
2) Generate and test a focused hypothesis about a corpus of texts, transforming an abstract interpretive question into a concrete quantitative task.  
3) Explain the theoretical complications of an empirical approach to literary narratives and acknowledge those complications as you construct your own DH project.  
4) Work collaboratively with your peers to both obtain and present your results, using the technical and interpretive strengths of each team member to enrich your final product.
Course Requirements

1) **Attendance and Participation.** This course is grounded in discussion and collaboration, so your consistent presence in class is crucial. If you know you need to be absent for a class session, you must notify me ahead of time in order to receive a make-up assignment. In-class activities may include large-group discussions, small-group projects, and individual writing assignments. 20% of final grade.

2) **Lab Visit and Report.** At least once over the course of the quarter, you must attend a meeting of the Stanford Literary Lab; I will keep you informed of relevant events. Take detailed notes during the meeting! You will need to write up a brief (approximately 500 words) report and analysis of the discussion. During the class session immediately following the lab meeting, you will summarize the meeting for your classmates and pose a few questions about the research you saw (no more than 10 minutes). If you and a classmate or two end up attending the same lab meeting, you may give a joint presentation, but you must each write your own report! 20% of final grade.

3) **Project Proposal (900-1200 words).** Before embarking on your final project, you and your team will write up a proposal articulating your research question, your selected corpus, your computational tools, and your expected results. A detailed assignment sheet will be distributed in class. 25% of final grade.

4) **Final Project Write-Up (1800-2400 words) and Presentation.** A detailed explanation of your final project, including your original hypothesis, research process, and analysis of results. You must situate your project in the context of a particular branch of literary study (e.g. genre theory, stylistics, narratology). During our final class session, you will present your preliminary results (approximately 15 minutes). A detailed assignment sheet will be distributed in class. 35% of final grade.

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 OAE is located at 563 Salvatierra Walk (phone: 723-1066, URL: [http://studentaffairs.stanford.edu/oaе](http://studentaffairs.stanford.edu/oaе)).

Honor Code

The Honor Code is the University’s statement on academic integrity written by students in 1921. It articulates University expectations of students and faculty in establishing and maintaining the highest standards in academic work:

The Honor Code is an undertaking of the students, individually and collectively:
1) 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.
2) 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.
3) 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.
4) 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.

Course Schedule

WEEK 1: What is a Text?
Monday
   Introduction: What is literature made of?

Wednesday
   Read before class: De Rose et al., “What is Text, Really?”
   Caton, “On the Term ‘Text’ in Digital Humanities"
   In class: experiment with Voyant

WEEK 2: Corpus vs. Canon
Monday
   Read before class: Cohen, “Reconstructing the Literary Field,” in The Sentimental
   Education of the Novel
   Moretti, “The Slaughterhouse of Literature”

Wednesday
   Read before class: Literary Lab Pamphlet 8, “Between Canon and Corpus: Six Perspectives
   on 20th-Century Novels”
   In class: building a corpus

WEEK 3: Style and Authorship
Monday
   Read before class: McEnery and Oakes, “Authorship Identification and Computational
   Stylometry”
   Hoover, “Corpus Linguistics, Stylometry, and the Styles of Henry James”
Wednesday
  Read before class: Literary Lab Pamphlet 5, “Style at the Scale of the Sentence”
  *In class: basic stylometry*

**WEEK 4: Where’s Waldo? Detecting Characters**

**Monday**
  Read during class: Twain, “The Celebrated Jumping Frog of Calaveras County”
  In-class exercise: how many characters are in this story? How would you design a program to identify *all* characters and *only* characters?

**Wednesday**
  Read before class: Vala et al., “On the Difficulty of Detecting Characters in Literary Texts”
  Bamman et al., “A Bayesian Mixed Effects Model of Literary Character”

**WEEK 5: Networks**

**Monday**
  Read before class: Literary Lab Pamphlet 2, “Network Theory, Plot Analysis”
  *In class: introduction to Gephi*

**Wednesday**
  Read before class: Elson et al., “Extracting Social Networks from Literary Fiction”
  Agarwal et al., “Social Network Analysis of *Alice in Wonderland*”

**WEEK 6: Beginning Your Project**

**Monday**
  Introduction to our shared corpus; metadata tutorial.

**Wednesday**
  Forming teams and generating research questions; wish list of computational tools.

**WEEK 7: Assembling Your Tools**

**Monday**
  Skills clinic, with guest instructor J.D. Porter. In-class work time.

**TUESDAY: Project Proposal Due by 5:00 p.m.!!**

**Wednesday**
  Project proposal conferences; in-class work time.

**WEEK 8: Operationalizing: Theory and Technique**

**Monday**
Read before class: Literary Lab Pamphlet 6, “‘Operationalizing’: Or, the Function of Measurement in Modern Literary Theory”

Wednesday
In-class exercise: with a partner, find a literary term or concept that could be operationalized. Develop a proposal for quantitatively measuring the concept, including a discussion of potential pitfalls or stumbling blocks.

WEEK 9: Ethics and Politics in DH
Monday
Read before class: McPherson, “Why are the Digital Humanities So White? Or, Thinking the Histories of Race and Computation”

Wednesday
Read before class: Allington et al., “Neoliberal Tools (And Archives): A Political History of Digital Humanities”
Cecire, “Theory and the Virtues of Digital Humanities”
Posner, “Money and Time”

WEEK 10: Wrap-Up and Final Conference
Monday
Collaborative writing workshop: how to divide the labor? In-class work time.

Wednesday
Final project presentations!

Friday of Week 11: Final project write-up due.