stanford university

F09-ENGLISH-253-01 : Literary Studies and the Digital Library

FALL 2009

Course Description

F09-ENGLISH-253-01 Exploration of a variety of ways of reading, interpreting, and understanding literature at the macro-scale, as an aggregate system. Consideration of theoretical issues, reading and discussion of landmark essays in the field, and an understanding of how digital libraries and literary corpora are inviting new types of literary research that frequently challenges conventional approaches.

Course Meetings

Section(s)

Name	Instructor	Day	Time	Location
Seminar Section(s)				
ENGLISH-253-01	Matthew Jockers	T,Th	1:15 PM - 3:05 PM	250-108

Course Syllabus

Overview:

Throughout this course we will investigate ways of reading, interpreting, and understanding literature at the macro or "corpus"-scale. Our "text" for the class will be a corpus of 1100 19th century British and American novels and our goal will be to produce a paper and submit our research for presentation at the 2010 meeting of the Alliance of Digital Humanities Organizations (ADHO) in London. (Should the paper be accepted, there are a variety of resources on campus and through ADHO that can assist or fully cover the travel costs).

Why:

Today's student of literature must be equally adept at gathering evidence from individual texts and accessing and mining digital text repositories. And "mining" here is the operative word. Literary scholars must learn to go beyond mere search. Today's students and scholars are adept at electronic search and comfortable searching digital collections for stray bits of evidence to fortify a hypothesis or support an argument. But the sheer amounts of literary "data" now available make search ineffectual as a means of thorough evidence gathering.

More interesting and exciting than the mere searching of digital archives is the ability to go beyond search and exploit computation to process and analyze the textual data these corpora contain. In practical terms, this means that the literary scholar must evolve, must embrace new approaches and new methodologies designed for accessing and leveraging the electronic texts that make up the 21st century digital library. This course provides a practical introduction to these methods in the context of a collaborative research project designed and executed by the students in the course.

Research and Collaboration:

In this course we will work collaboratively. Collaborative scholarship is common in the sciences but somewhat foreign to the humanities. That said, in the nascent discipline of "Digital Humanities" (or "Humanities Computing") where projects are often ambitions and require teams with diverse skill sets, collaboration is becoming a norm rather than an exception. See for example Lisa Spiro's recent blog posts about collaboration (Collaborative Authorship in the Humanities and Examples of Collaborative Digital Humanities Projects).

Assessment:

Research Project:Most of the course will revolve around the construction and execution of a collaborative research project. Students will collectively form a research question and design and implement an approach to investigate the problem. There will be various stages in the evolution of the project (see "schedule" below). The critical final deliverable will be a paper proposal which we will submit to the ADHO conference hosts for presentation. Examples of successful proposals (including one submitted by students in my course last fall) will be provided as models. The project will account for 40% of your grade.

Exercises: A series of exercises have been prepared in order to give you hands on experience with a variety of computational methodologies for analyzing text. No prior programming experience is required or assumed. The exercises are gentle and designed to be enjoyable. Each exercise has a required component and an "extra credit" challenge question. The exercises account for 30% of your grade. (answers here)

Methodology Reports: Each student will prepare a 15 minute presentation a lead a followup discussion on a instructor-assigned topic. Typically this will involve reading and summarizing for the class some essay related to our work in the course. The presentation will account for 10% of your grade.

Collaboration: A collaborative enterprise can not succeed without commitment and participation. Attendance, punctuality, and participation in class will be expected and will constitute 10% of your grade.

Group Evaluation: Each student will write a 250-500 word evaluation of their team. This should be a sort of meta-analysis of the group dynamic: what worked well, how did the group self-organize, how was the work distributed, etc. The evaluation will account for 10% of your grade.

Schedule

1. T. 9/22 - Introduction to Course: Activate your cgi space.

   Th. 9/24 History of Literary and Linguistic Computing: Readings: Horowitz, "Visualizing Big Data: Bar Charts for Words"; Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete"; Hockey, "History of Humanities Computing"

2. T. 9/29 The Digital Library of Today and Tomorrow: Guest Lecture by Glen Worthey. Readings: Kelly, "Scan this Book"; Grafton, "Future Reading: Digitization and its Discontents"; Toobin, " Google's Moon Shoot"; Hirschorn,"The Hapless Seed"; Crane, "What Do You Do with a Million Books"; Hafner, "History Digitized (and Abridged)"

   Th. 10/1 Literary Studies: Readings: Rommel, "Literary Studies" (Kraus); Hoover, "Quantitative Analysis and Literary Studies" (Sajor); Cummings, "The Text Encoding Initiative and the Study of Literature" (Law)

3. T. 10/6 Reading: Essays by Friedlander and by Oard in the Clir Report on the Next Generation of Digital Humanities Scholarship; Ramsay, "Algorithmic Criticism" (Sudhof); also Exercise One Due

   Th. 10/8 Making the Most of Metadata. Readings: Moretti, "Style inc. Reflections on 7,000 Titles" (Heuser) See also (Charts); Jockers, "Selection from *Draft* Manuscript titled 'Beyond Search: Literary Studies and the Digital Library"

4. T. 10/13 Exercise Two Due

   Th. 10/15 Corpus Building and Text Encoding: Readings: Renear, "Text Encoding" (Paymer); Willet, "Electronic Texts: Audiences and Purposes" (Truxaw); Ide, "Preparation and Analysis of Linguistic Corpora" (Zhu)

5. T. 10/20 Exercise Three Due

   TH. 10/22 Text Processing: Readings: Burrows, "Textual Analysis" (Tevel); Hajic, " Linguistics Meets Exact Sciences" (Lewis);

6. T. 10/27 Exercise Four Due

   Th. 10/29 Guest Visit from Jenny Loomis. Project Abstracts Due.

7. T. 11/3 Exercise Five Due

   Th. 11/5 Authorship Attributions and Stylistic Analysis: Readings: Craig, "Stylistic Analysis and Authorship Studies" (Kharputly); Jockers et. al, "Reassessing Authorship in the Book of Mormon Using Delta and Nearest Shrunken Centroid Classification"; Jockers and Witten, "A Comparative Study of Machine Learning Methods for Authorship Attribution"; Crandell, et. al. "The Sound and the Fury Final Project"

8. T. 11/10 Exercise Six Due

   Th. 11/12 Got Theory? Readings: Tynjanov, "On Literary Evolution" (Le-Khac); Ejxenbaum, "The Theory of

the Formal Method" (VanArendonk); Shklovsky, "Art as Device" (Blevins)

9. T. 11/17

   Th. 11/19

10. T. 12/1

    Th. 12/3

go to CourseWork site ▶

For course site members only