

Sp12-ENGLISH-162-01 : Critical Methods: Introduction to Digital Humanities

SPRING 2012

Course Description

Sp12-ENGLISH-162-01 Introduction to the different intellectual models which help us explain and interpret literary texts, genres, and movements.

Digital texts and digital libraries offer us new opportunities for searching and accessing literary material. But more interesting and exciting than the mere searching of digital texts is the ability to leverage computation in order to process and analyze textual data, to provide new methods for reading, analyzing, and understanding literature.

This course provides an introduction to the field of humanities computing with a special emphasis on literary text-analysis. Students learn about the preparation and processing of digital texts while exploring literary methods which help us explain and interpret literary texts, genres, and movements. The course includes units dealing with "stylometry" (computer based stylistic analysis), authorship attribution, gender detection, text encoding, and the visualization of literary information using such open source tools as R and Gephi.

Throughout the course we consider the theoretical issues associated with employing quantitative methodologies in a traditionally qualitative discipline; we read and discuss landmark essays in the field; and we end with an informed discussion of how digital libraries and computation are taking literary scholarship "beyond the book." Students will develop basic coding skills in an environment in which understanding literature is the only prerequisite. No programming experience is required; students will develop fluency in XML and R through exercises and work on a collaborative text-analysis project.

Course Meetings

Section(s)

Name	Teaching Staff	Dates	Day	Time	Location
Lecture Section(s)					
ENGLISH-162-01	Matthew Jockers	04/02/12 - 06/06/12	T,Th	1:15 PM - 3:05 PM	200-219

Course Syllabus

Digital Humanities: Beyond the Book, 2012

All Digital Humanities projects have what might be described as a "life cycle" or process of maturation from idea to outcome. To the extent that it is possible, this course is organized along the lines of a real-world digital humanities project. In fact, the projects we will complete in this class are real research projects in every sense of the word. The only real difference here is that the projects will be completed on an accelerated timeline.

A note about digital source material: All of our research in this class will be based on a computational analysis of works written by Virginia Woolf. Early in the term, Professor Alice Staveley will be visiting our seminar to talk about some of her ideas about this corpus and the kinds of questions that scholars of Woolf might like to explore (or have explored for them :-). Professor Staveley will be available at other times during the term to consult with us as the projects mature.

This course has five major components or "phases," as follows:

1. Preparation and Practice. In weeks 1-7, you will become familiar with critical methods and materials of digital humanities research. Through class readings and participation in class discussions you will learn about major trends in the field. Through completion of programming exercises you will practice various forms of text-analysis that are possible using the R programming language.

2. Project Proposal. On or before May 1, each research group* will submit a research proposal following the model of a Level I Digital Humanities Start-Up Grant as specified in Section IV numbers 1-4 of the NEH web page: (<http://www.neh.gov/grants/guidelines/digitalhumanitiesstartup.html>). This proposal will account for 10% of your overall grade (20% of the Project Grade).

*Research group composition will be determined during our first meeting of the second week.

3. Exploration. Weeks five, six and seven will be spent on data processing and exploration. This will be the most intense, potentially unfamiliar, and exciting time of the term. During this exploratory/experimental phase you will be doing some "fishing": pursuing your group's ideas, hunches, and hypotheses in an effort to extract from the corpus what will become the meaningful data points in support of your larger argument. The coding skills you develop in the first five weeks will be critical here. During this phase, you should save and comment all of your code. This code will be submitted as an appendix to the final essay and will account for 10% of your final grade.

4. Analysis and Conclusions. Weeks eight and nine and ten will be devoted to analysis and synthesis of the data derived during the exploration phase. This is the period in which you bring order to the chaos of your experimentation. You will move, as Heuser and LeKhac write, from "signal" to "concept" and then from concept to interpretation and argument. Here you will form the results of your enquiry into a collaboratively written scholarly essay about Woolf's corpus. Your objective will be to show us something new! The essay will account for 1/2 of the project grade or 20% of your final grade. The essay is due in class on June 5.

5. Project Report: In place of a final exam, your group will submit a concise ~1000 word report to the agency that funded your project proposal. This report should include a 150-200-word executive summary followed by an honest report of the outcomes of your project and how the final project evolved in relation to the plan articulated in the original proposal. The report should include a summary of each participant's contribution to the project. The report will account for 1/4 of the project grade or 10% of your final grade. (Due June 13).

Required Materials

A computer, preferably a laptop you can bring to class.

[R \(open-source programing language and console qui\)](#)

A code editing application such as Notepad++, TextWrangler, Jedit, etc. Here is a link to a bunch of open source editors that I know work well with R. <http://www.sciviews.org/rqui/projects/Editors.html>

Assessment

- Exercises: 40%
- Group Project: Proposal 10% - Code 10% - Essay 20% - Report 10%
- Participation (including discussion of readings): 10%

Schedule

Week I

Tuesday (4/3)

- Download and Install R. [Instructions for Downloading R](#)
- Download and Install the Exercise Corpus. [Instructions for Downloading the corpus](#)

Thursday (4/5)

- [The Digital Humanities and Humanities Computing: An Introduction](#)
- [Hockey, The History of Humanities Computing](#)
- [Burnard, Lou. "Is Humanities Computing an Academic Discipline? or, Why Humanities Computing Matters." 1999. Web. 11 Nov 2009.](#)
- [Busa, Roberto. "The Annals of Humanities Computing: The Index Thomisticus." Computers and the Humanities. 14 \(1980\): 83-90](#)
- [Literary Studies](#)
- [Potter, Literary Criticism and Literary Computing. . .](#)

Week 2

Tuesday (4/10)

DUE TODAY: [Exercise 1, Word Frequencies Tabulation](#)

Project Groups Determined in Class Today.

- [Textual Analysis](#)

- [Siemans, A new Computer-Assisted Literary Criticism](#)
- [Rockwell, What is text analysis, really?](#)
- [Algorithmic Criticism](#)
- [Sinclair Computer-Assisted Reading: Reconceiving Text Analysis](#)

Thursday (4/12)

- [Quantitative Analysis and Literary Studies](#)
- [Hoover, Frequent Word Sequences. . .](#)
- [Linguistics Meets Exact Sciences](#)
- [Stylistic Analysis and Authorship Studies](#)
- [Burrows Not Unless You Ask. . .](#)

Week 3

Tuesday (4/17)

DUE TODAY: Exercise 2, Accessing and Comparing Word Frequency Data

- [Quantitative Formalism: an Experiment](#)

Thursday (4/19)

DUE TODAY: Exercise 3, Word frequency over Novelistic Time

Alice Staveley visits class to talk about the Woolf Project.

Week 4

Tuesday (4/24)

DUE TODAY: Exercise 4, Correlation

- [Reflections on Seven Thousand Titles](#)

Thursday (4/26)

- [Koppel, Automatically Categorizing Written Texts by Author Gender](#)

Week 5

Tuesday (5/1)

DUE TODAY: Exercise 5, Two Types of Vocabulary Richness and Another Use of Correlation

Thursday (5/3)

Week 6

Tuesday (5/8)--Guest Lecture with Aaron Stanton, CEO of [Booklamp.com](#)

DUE TODAY: Exercise 6, A KWIC List Application

Thursday (5/10)--Guest Lecture with Ryan Heuser of the Stanford Literary Lab

- [Heuser and Le-Khac, Learning to Read Data](#)

Data Exploration

Week 7

Tuesday (5/15)

DUE TODAY: Exercise 7: KWIC List

Data Exploration:

Thursday (5/17)

Data Exploration:

Week 8

Tuesday (5/22)

DUE TODAY: Exercise 8

Data Exploration:

Thursday (5/24)--Guest Lecture with Glen Worthey, Stanford's Digital Humanities Librarian.

Reading:

Tynianov: Literary Evolution

Worthey: Russian Formalism and DH

Shklovsky: Art as Device

Eikhenbaum: Theory of the Formal Method (SKIM this one if you want! It's long)

Data Analysis and Synthesis

Week 9

Tuesday (5/29)

Data Analysis and Synthesis

DUE TODAY: Exercise 9

Thursday (5/31)--Guest Lecture with Jodie Archer

Data Analysis and Synthesis

Week 10

Tuesday (6/5)

Data Analysis and Synthesis

Thursday (6/7)

Data Analysis and Synthesis

[go to CourseWork site](#) 

For course site members only