stanford university

F08-ENGLISH-153H-01/HUMNTIES-198J-01: Digital Humanities: Beyond the Book

FALL 2008

Course Description

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.

"Digital Humanites: Beyond the Book" 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 questions. 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 tools as Many Eyes, Google Earth, Excel, and custom tools we develop in the class.

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 PHP through exercises and work on a collaborative text-analysis project.

For more information go to this website:

Course Meeting(s)

Course Syllabus

Description:

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.

"Digital Humanites: Beyond the Book" 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 questions. 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 tools as Many Eyes, Google Earth, Excel, and custom tools we develop in the class.

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 PHP through exercises and work on a collaborative text-analysis project.

Materials

Companion to Digital Humanities

oXygen XML Editor (\$48.00) via Oxygen Web Site

Assessment

- 40% Exercises
- 40% Group Project (pdf)
- 10% Participation
- 10% Group Evaluation

Schedule

Week I, Tuesday: Course Introduction

• Activate Your Personal CGI Service. (Let me know asap if you have any trouble.)

Week I, Thursday: Foundations

Class Discussion: Please have these read before class

- Humanism
- Humanities
- What are the Humanities
- Humanities
- Hamming, R. W. "The Unreasonable Effectiveness of Mathematics." (requires sunet authentication)
- Bush, Vannevar. "As We May Think"
- Laue, Andrea, "How the Computer Works"

Week II, Tuesday: Humanities Computing Overview

Class Discussion: Please have these read before class

- Busa, Roberto. "Perspectives on the Digital Humanities"
- Susan Schreibman, Ray Siemens, and John Unsworth. "The Digital Humanities and Humanities Computing: An Introduction"
- Hockey, Susan. "The History of Humanities Computing"
- McCarthy, Willard. "What is Humanities Computing"
- Rommel, Thomas. "Literary Studies"

Week II, Thursday

- Exercise One Due
- Download "<u>Directory Indexer</u>" (a file called "index.php"--don't change the name) and put it into your digHum directory (created as part of Exercise one). This file will create a clickable index of all the sub directories and files in your DigHum folder

Week III, Tuesday

Lecture

Week III, Thursday

• Exercise Two Due: Simple HTML and PHP

Week IV, Tuesday: Authorship Attribution

Class Discussion: Please have these read before class

- Craig, Hugh. "Stylistic Analysis and Authorship Studies"
- Burrows, John. "Not Unless You Ask Nicely: The Interpretive Nexus Between Analysis and Interpretation
- Burrows, John. "Who Wrote Shamela: Verifying the Authorship of a Parodic Text"
- Hoover, David. "Frequent Word Sequences and Statistical Stylistics"
- Koppel, Moshe (et. al). "Automatically Categorizing Written Texts by Author Gender"
- Merriam, Thomas. "Linguistic Computing in the Shadow of Postmodernism"

Also this week:

- Group Assignements:
- Group I TBD
- Group II TBD

Week IV, Thursday:

• Exercise Three Due: Text Analysis and Authorship Attribution

Week V, Tuesday: Text Encoding

Class Discussion: Please have these read before class

- Renear, Allen. "Text Encoding"
- Willett, Perry. "Electronic Texts: Audiences and Purposes"
- McGann, Jerome. "Marking Texts of Many Dimensions"
- Popham, Michael. "Text Encoding, Analysis, and Retrieval"

- A Gentle Introduction to the TEI markup language
- Huitfeldt, Claus. "Scholarly Text Processing and Future Markup Systems"
- Eggart, Paul. Text-Encoding, Theories of the Text, and the 'Work-Site'

Week V, Thursday

Exercise Four Due: Introduction to TEI

Week VI, Tuesday: Text Transformation

- Lecture
- Learning XSLT Not Required reading, but an excellent resource available to all Stanford Students online.

Week VI, Thursday

Exercise Five Due: XSLT

Week VII, Tuesday: DOM XML

• Lecture

Week VII, Thursday

Exercise Six Due: Parsing XML with php

Week VIII, Tuesday: Basic Text Analysis

Class Discussion: Please have these read before class

- Rockwell, Geoffrey. What Is Text Analysis, Really?
- Sinclair, Stefan. Computer-Assisted Reading: Reconceiving Text Analysis
- Siemans, Ray. A New Computer-Assisted Literary Criticism?
- Potter, Rosanne. Literary Criticism and Literary Computing: The Difficulties of Synthesis
- Burrows, John. Textual Analysis
- · Rockwell, Geoffrey. Mimes and Mermaids.pdf
- Programming PHP Not Required reading, but an excellent resource available to all Stanford Students online.

Week VIII, Thursday

Exercise Seven Due

Week IX, Tuesday

Lecture

Week IX, Thursday

Exercise Eight Due: Text Searching with php and XML

Week X: Thanksgiving Break

Week XI, Tuesday: The Future of Humanities Computing

Class Discussion: Please have these read before class

• Hockey, Susan. Living with Google: Perspectives on Humanities COmputing and Digital Libraries

Week XI, Thursday: Final Presentations



for course site members only