PLAI 600: Interdisciplinary Theories and Methods
LLCU 606 : Topics in Literature

Reading Digital Texts

Professor S. Sinclair
Mondays 2-4pm, room 465, 688 Sherbrooke (Fall 2012)
Office Hours (#341, 688 Sherbrooke): Mondays 1-2pm or by appointment

Description

Digital texts are composed of discrete units of information that have the virtue of being infinitely malleable and reconfigurable, allowing new practices for searching, filtering, comparing, annotating, measuring, representing and understanding texts. From single works to virtual libraries, from canonical classics to contemporary social media, digital texts can provide rich fodder for interpretive practices in the digital humanities. This course will provide students with theoretical and practical foundations for working with a variety of digital texts.

Learning Outcomes

By the end of this course students should be able to

- compile and prepare for analysis a variety of types of digital texts
- use text analysis and statistical software to study digital texts
- produce visualizations from analytic and statistical data
- think algorithmically about digital texts
- navigate between close and distant reading practices
- describe and demonstrate the value and limitations of digital methodologies

Course Content

This is essentially a course on computer-assisted text analysis and text mining from humanistic perspectives. Considerable time and effort will be devoted to learning new digital methodologies. Students need not have any technical background, but students do need to be thoroughly committed to learning new skills that will require abundant perseverance and resourcefulness. The primary purposes of this course are to learn to think algorithmically about digital texts and to explore what interpretive and hermeneutic practices are enabled by digital methodologies (learning technical skill are a means to those ends). The course is organized into a set of modules whose duration will be determined by the students interests and needs:

- First Encounters with Digital Text & Tools
    - what digital texts are out there?
    - what digital tools are out there (Voyant Tools, Monk Project, Paper Machines, Wordseer, etc.)?
    - what kinds of things are people doing with digital texts & tools?
    - what is a digital humanities approach to text analysis and text mining?

- Close Reading and Microanalysis with R
  - loading a text
  - generating and plotting frequencies
  - graphing distribution
  - correlating & comparing data
  - measuring vocabulary richness
  - generating concordances
  - working with XML
- Preparing a Corpus
  - data sources & formats
  - web harvesting
  - cleaning & converting
- Distant Reading and Macroanalysis with R and Mallet
  - loading a corpus
  - loops and functions
  - measuring distance between texts
  - topic clustering
- Text & Data Representations
  - exporting, converting, and importing data
  - visualization: spreadsheet applications, specialized applications, web-based solutions

Instructional Method

This course is a graduate seminar that is a mix of discussion and hands-on workshop.

Course Materials

There are no materials to purchase for this course, all materials (weekly readings, tutorials, exercises, etc.) will be made available in digital format through the course website. The primary text that will be used is *Micro and Macro Analysis Using R* by Matthew Jockers (forthcoming) – this will be complemented by a selection of additional relevant articles. Students need to have a laptop computer with them during every class. Software packages include RStudio and Gephi.

Assignments & Evaluations

Component	Weight	Notes
weekly blog posts:	20%	weekly reflections on technical aspects are essential
assignments	40%	one assignment for each of the four modules
final project	30%	a final, larger, collaborative project
participation & collaboration	10%	active and insightful participation and collaboration