

Cultural Analytics: The Computational Study of Culture

LLCU 609 – Winter 2015
T 2:35 - 4:25, Rm. 325 688 Sherbrooke

Professor Andrew Piper
Office: Rm 483, 688 Sherbrooke
Phone: 514-398-4400 x094504
Email: andrew.piper@mcgill.ca
Office Hours: T 12 – 1 pm

Course Description

This course will serve as a critical introduction into the new tools and techniques that are being developed to study literature and culture at a vastly greater scale. Questions we will be asking will include: How does the ability to analyze hundreds of thousands of documents give us new insights into the history of different cultural forms? How can we import key cultural theoretical categories like discourse, power, belief, narration, reception, signification and authorship into the computational study of culture? And how might computation impact how we theorize culture? The aim of the seminar is to engage critically with existing methods and practices in the computational study of culture, including stylistics, sentiment analysis, topic modeling, social network analysis and geospatial tools, as well as offer a creative platform through which we can collectively develop new ideas and hypotheses about studying culture at large scale. No prior programming experience or knowledge of statistics is required, just open-mindedness and a willingness to learn.

Weekly Assignments

All readings are available on MyCourses.

Wk. 1 01.06 *Course Introduction*

Wk. 2 01.13 *What is Culture?*

- Longhurst et al., *Introducing Cultural Studies*, pp. 1-57
- Smith/Riley, *Cultural Theory*, “Introduction,” pp. 1-5 and “Structuralism and the Semiotic Analysis of Culture,” pp. 92-110.
- Michel, “Quantitative Analysis of Culture Using Millions of Digitized Books.”

Wk. 3 01.20 *Style and Vocabulary Richness*

- Burrows, “Delta: A Measure of Stylistic Difference.”
- Lancashire, “Vocabulary Change in Agatha Christie.”
- Matt Daniels, “The Largest Vocabulary in Hip Hop”:
<http://rappers.mdaniels.com.s3-website-us-east-1.amazonaws.com>

- Wk. 4** 01.27 *Diction*
- Underwood/Sellers, “The Emergence of Literary Diction”:
<http://journalofdigitalhumanities.org/1-2/the-emergence-of-literary-diction-by-ted-underwood-and-jordan-sellers/>
 - Heuser/Le-Khac, “A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method.” Literay Lab Pamphlet 4 (May 2012):
<http://litlab.stanford.edu/LiteraryLabPamphlet4.pdf>.
 - *Extra Credit*: Adam Kilgarriff, “Comparing Corpora.”
- Wk. 5** 02.03 *Sentiment Analysis (Belief 1)*
- <http://hedonometer.org>
 - Liu, “Sentiment Analysis and Subjectivity.”
 - Jockers, “A Novel Method for Detecting Plot.”
<http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/>
- Wk. 6** 02.10 *Topic Modeling 1*
- Mohr, “Topic Models: What they are and why they matter.”
 - Schmidt, “Words Alone: Dismantling Topic Models.”
<http://journalofdigitalhumanities.org/2-1/words-alone-by-benjamin-m-schmidt/>
 - Underwood, “Topic Modeling Made Just Simple Enough.”
<http://tedunderwood.com/2012/04/07/topic-modeling-made-just-simple-enough/>
 - DiMaggio et al, “Exploiting affinities between topic modeling and the sociological perspective on culture.”
 - Jockers/Mimno, “Significant Themes in 19C Literature.”
- Wk. 7** 02.17 *Topic Modeling 2*
- McFarland et al, “Differentiating language usage through topic models.”
 - Marshall, “Defining Population Problems: Using topic models for cross-national comparison of disciplinary developments.”
 - Underwood/Goldstone, “The Quiet Transformations of Literary Studies.”
- Wk. 8** 02.24 *Social Network Analysis*
- Kieran Healy, “Using Metadata to find Paul Revere.”
<http://kieranhealy.org/blog/archives/2013/06/09/using-metadata-to-find-paul-revere/>
 - Newmann, Networks, “Introduction,” & Chap. 4.
 - So/Long, “Network Analysis and the Sociology of Modernism.”
 - Newmann, “Co-Authorship Networks and Patterns of Scientific Collaboration.” http://www.pnas.org/content/101/suppl_1/5200.full

03.03

Study Break

Wk. 9 03.10

Character Networks

- Elson, "Extracting social networks from literary fiction."
- Piper, "Communities of Detection: Social Network Analysis and Detective Fiction."
- Bamman/Underwood/Smith, "A Bayesian Mixed Effects Model of Literary Character."

Wk.10 03.17

Narration and World View (Belief 2)

- Genette, *Narrative Discourse*, Chap. 5
- Underwood, "Gender, Genre, and Point of View":
<http://tedunderwood.com/2013/09/22/genre-gender-and-point-of-view/>
- Piper, "Novel Devotions: Conversional Reading, Computational Modeling, and the Modern Novel."
- Schmidt, "Fundamental Plot Arcs on TV":
<http://sappingattention.blogspot.com/2014/12/fundamental-plot-arcs-seen-through.html>

Wk.11 03.24

Circulation

- Genette, *Palimpsests*, pp. 1-13
- Hall, "Encoding/Decoding."
- Cordell, "Infectious Texts."
- *Commonplace Cultures*:
http://commonplacecultures.org/?page_id=99

Wk.12 03.31

Geospatial Analysis

- Wilkens, "The Geographic Imagination of Civil War Era American Fiction."
- *Spatial History @ Stanford*:
<http://web.stanford.edu/group/spatialhistory/cgi-bin/site/index.php>
- Anne Knowles: <http://empirestudies.com/2013/09/20/what-could-lee-see-at-gettysburg/>

Wk.13 04.07

Visual Culture

- <http://selfiecity.net>
- <http://phototrails.net>
- Piper, "Footnote Detection": <http://txtlab.org/?p=395>

Academic Integrity

McGill University values academic integrity. Therefore all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see <http://www.mcgill.ca/integrity/> for more information).

Course Requirements

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded.

Class Participation	20%
Presentation	15%
Final Paper	65%

Class Participation. You are expected to attend every class and **actively** participate in class discussions with observations and questions derived from close and thoughtful reading of each weeks' texts. Passive consumption of information is not welcome. Our aim is to engage critically with existing quantitative studies of culture and creatively with the new possibilities they initiate.

Presentation. The presentation will take the form of a hypothetical research study design. Given all the resources in the world, what would you want to study computationally and how? What are the steps necessary to undertake such a project, from hypothesis creation to data collection to data preparation to data analysis, and where do you see potential problems? Each presentation will consist of a 10-minute description of a proposed research study for which we will collectively provide feedback. The goal is to better understand the challenges and compromises involved in moving from a theory of culture to the implementation of the study of culture.

Final Paper. For the final paper you will write a critical review of a particular study that we have *not* read in class. This will involve, first, locating a significant paper in a particular field and, second, engaging with all the steps of the research process as they are presented in the paper. The point of this exercise is to have you learn to reflect on the choices that are made at each step of a study, to understand why those choices were made, explain your concerns as clearly and concisely as possible, and then offer possible alternative pathways or interpretations of the data, drawing on other research in the field to support your views. The paper should be between 12-15 pages in length and is due on **April 14, 2015**.

Late papers will lose a half-grade for every class late. Students who receive a grade of D,F, or J will not be allowed to do supplemental work. All papers will be submitted to the text-matching software per university policy. Three or more missed classes will result in a lowering of the student's overall grade. According to Senate regulations, instructors are not permitted to make special arrangements for final exams. Please consult the Calendar, section 4.7.2.1, General University Information and Regulations at www.mcgill.ca. In the event of extraordinary circumstances beyond the University's control, the content and/or evaluation scheme in this course is subject to change. © Instructor generated course materials (e.g., handouts, notes, summaries, exam questions, etc.) are protected by law and may not be copied or distributed in any form or in any medium without explicit permission of the instructor. Note that infringements of copyright can be subject to follow up by the University under the Code of Student Conduct and Disciplinary Procedures.