

## EMPIRICAL METHOD IN LITERARY STUDIES

English 505.401  
Fall 2013  
Thurs 12:00-2:50  
Kislak Center 629, Van Pelt

Jim English  
Office: FBH 311  
Hours: Tues 2:30-3:30 /Thurs 3:30-4:30  
jenglish<at>sas.upenn.edu

The use of empirical methods – involving collection and analysis of data – is not new in literary studies, but it has always been a marginal practice in a discipline dominated by qualitative forms of historical and hermeneutical criticism. Lately, however, scholars exploring the potential of digital tools for literary analysis as well as those proposing new articulations of literary studies with sociology have reopened the question of empirical method. In this class, we will look at some emergent quantitative approaches together with the theoretical and political debates they have occasioned. We will consider the core concept of “data” and the ontology of datasets as these pertain to literature; the kinds of research project that can or cannot be assisted by surveys, statistics, and measurements; and the social and institutional stakes of embracing a computational paradigm in the humanities. We will all learn a little bit of computer programming, using the open source programming language R for statistical analysis and graphics. Programming tutorials and other assistance will be available throughout the term in the Vitale II media lab.

This is an introductory-level graduate class, open to highly advanced undergraduate majors by permission. Assigned reading will be fairly heavy at 150-200 pages per week and everyone, including auditors, is expected to complete the readings prior to class meetings. Other required work will include short essays and thought experiments, simple computer programming tutorials and problem sets, data-based projects, and oral presentations. As this is a 500-level course, no full-length research paper will be assigned and no particular background knowledge is expected.

We will be using a Canvas site for this class. If you are unfamiliar with Canvas and would like a tutorial, let me know and I’ll make arrangements with the Library. Nearly all assigned readings are available in pdf format from the Electronic Reserves tab on the Canvas site. Where a reading has been excerpted from a book, the book is on reserve at the Rosengarten desk in Van Pelt lower level. Please purchase a print or Kindle copy of Franco Moretti’s *Distant Reading* (Verso 2013).

Much of our class time will be spent on presentations, debates, small-group activities, project planning, and problem solving. Several of our class meetings will be workshops or tutorials led by guest scholars.

## SCHEDULE

Aug 28 Introduction: Objectives and Expectations

Workshop 1: Keyword definitions

Workshop 2: Brainstorming in-class activities and exercises

Tech skills: how we will learn to program using R

### **I. The Data of Literary Reception: Readers, Taste, Commerce**

Sept 4 Taste, Class, Gender, Habitus

Pierre Bourdieu, "Introduction to the First Edition" and "Part II: The Economy of Practices," from *Distinction* (1984)

Janice Radway, "Introduction," Chapters 2 and 3, "Conclusion," Appendixes," from *Reading the Romance* (1991)

Sept 11 National Readerships: India, Nigeria, Australia

GUEST SCHOLAR: Priya Joshi

Priya Joshi, "Quantitative Method and Literary History" (2002)

Joshi, "Consuming Fiction" from *Another Country* (2002)

Robert Darnton, "Reading, Writing, and Publishing in 18<sup>th</sup> Century France" (1971)

Wendy Griswold, "The Readers" and Appendixes from *Bearing Witness: Readers, Writers and the Novel in Nigeria* (2000); "Number Magic in Nigeria" (2002)

John Frow and Tony Bennett, "Reading by Numbers," from *Accounting For Tastes* (1999)

Frow, "On Midlevel Concepts" (2010)

Sept 18 Literary Commerce: Readers as Consumers

William St Clair, "Reading Constituencies" from *The Reading Nation in the Romantic Period* (2006)

St Clair, "The Political Economy of Reading" (2005/2012)

John B. Thompson, "The Polarization of the Field" from *Merchants of Culture* (2010)

Tim Aubry, "Aghanistan Meets the Amazon" (2009)

Ed Finn, "Becoming Yourself: The Afterlife of Reception" (2011)

### **II. Linguistic Data: Algorithmic Approaches to Content, Meaning, and Style**

Sept 25 Stylistics and its Critics

Stanley Fish, "What is Stylistics and Why Are They Saying Such Terrible Things About It?" and "What is Stylistics . . . About It, Part II" (1979)

Douglas Biber, "Corpus Linguistics and the Study of Literature: Back to the Future?" (2011)  
Michaela Mahlberg, "Corpus Stylistics" and "Character Speech" from *Corpus Stylistics and Dickens's Fiction* (2013)  
Allison et al., "Style at the Scale of the Sentence," *Stanford Literary Lab Pamphlet 5* (2013)

Oct 2            Machine Reading

Daniel Rosenberg, "Data Before the Fact," in Gitelman, ed. *Raw Data is an Oxymoron* (2013)  
Stephen Ramsay, "An Algorithmic Criticism," and "Potential Readings" from *Reading Machines: Toward an Algorithmic Criticism* (2011)  
Matthew Jockers, "Macroanalysis" and "Style" from *Macroanalysis: Digital Methods & Literary History* (2013)

Oct 9            No Class, Fall Break

Oct 16          Topic Modeling

DHF Workshop 12:00-1:30  
GUEST SCHOLAR: Andrew Goldstone

Petko Bogdanov and John Mohr, "Introduction—Topic Models: What They Are and Why They Matter" (2013)  
Andrew Goldstone, "[Social Science and Profanity at DH2014](#)" (2014)  
Andrew Goldstone and Ted Underwood, "The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us" (forthcoming); [Quiet Transformations website](#)  
Matthew Jockers and David Mimno, "Significant Themes in 19<sup>th</sup>-Century Literature" (2013)  
The Programming Historian "[Getting Started with Topic Modeling and Mallet](#)"

Oct 23          Workshop: Text Mining with R

GUEST SCHOLAR: Katie Rawson

Assignments TBA

### III. Big Literary Data: Maps, Systems, and the Circulation of Forms

Oct 30          Distant Reading and its Critics

Franco Moretti, "Conjectures on World Literature," "The Slaughterhouse of Literature," "More Conjectures," "Evolution, World-Systems, *Weltliteratur*," "The Novel: History and Theory," from *Distant Reading*

Andrew Milner, "Literature, Culture, and Society" pp. 116-140

Katherine Bode "Literary Studies in the Digital Age," from *Reading By Numbers* (2012)

Eric Hayot, "Scale, Data, and World Literature" (slideshow)

Polemics:

Gayatri Chakravorty Spivak, "World Systems and the Creole" (2006)

David Golumbia, "Death of a Discipline"

Katie Trumpener, "Paratext and Genre System: A Response to Franco Moretti"

Moretti, "Relatively Blunt"

Nov 6                      Circulation and Mapping

DHF Workshop 12:00-1:30

GUEST SCHOLAR: Ryan Cordell

Ryan Cordell, "Uncovering Patterns of Reprinting in Antebellum American Newspapers" (forthcoming)

Mapping the Republic of Letters: Case Studies

Caroline Winterer, "Where is America in the Republic of Letters?" (2012)

#### IV. Visualizing Data: Graphesis and Knowledge Production

Nov 13                      Visualization and Interpretation

Johanna Drucker, "Humanities Approaches to Graphical Display" (2011)

Stéfan Sinclair, Stan Ruecker, and Milena Radzikowska, "[Information Visualization for Humanities Scholars](#)" online at MLACCommons (2013)

Nov 20                      Character Networks and Character Space

UWGB Dracula Character and Text Network

Franco Moretti, "Network Theory, Plot Analysis" (2011)

Moretti, "Operationalizing, or the Function of Measurement in Modern Literary Theory," Stanford Literary Lab Pamphlet 6 (2013)

Graham Sack, "Character Networks for Narrative Generation: Structure Balance Theory and the Emergence of Proto-Narratives" (2013)

Sack, "Simulating Plot: Towards a Generative Model of Narrative Structure" (2011)

Nov 27                      No Class, Thanksgiving

Dec 4                        Final Class: Lightning-Round Presentations