

SOC 201B: Sociology of Big Data, Privacy and Surveillance

Term: Summer 2016

Meeting schedule: Tue/Thu 10:20-12:20 in ART 003

Credit Hours: 5

Instructor: Nina Cesare (ninac2@uw.edu)

Office Hours: Wednesday 1:00-3:00 in Savery 250 (or by appointment)

Course Description:

The past decade has seen an emergence of tools that provide insight into the workings of the social world, and social scientists have begun to think critically about how to systematically capture and analyze these data. The availability of digital traces from cell phones, social media, and other sources have opened up a “new era of social research,” and it is our goal in this course to understand the methodological and ethical landscape of this growing social science subfield.

The first few weeks of the quarter will focus first on the definition of “big data,” the breadth and depth of these data, and the significance of these data within the social sciences. The latter part of the quarter will highlight issues related to privacy and surveillance that the availability of big data raise and will address the ethical implications of using these data for social science research. The majority of assignments for this course will be written reflections that synthesize and build upon material covered in lecture. This class will also involve at least one hands-on lab activity through which students may view and analyze segments of digital trace data. The course will conclude with an independent research paper on a topic of your choosing.

Learning Goals:

After taking this course, you should be able to do the following:

- Be able to define, in your own words, what the term “big data” means in the context of social science research.
- Think critically about how to address the benefits and limitations of big data in social science research.
- Understand the privacy-related and ethical implications of conducting social science research using big data.

Course Expectations:

Everyone is expected to help create an environment conducive to learning in which your fellow students feel comfortable sharing their ideas. To achieve this, I expect you to do the following:

- Show up for class (the easy part, but also half the battle).
- Complete assignments and readings on time.
- Engage in fair and open discussion with your fellow classmates. Let them speak and carefully consider your response. If you disagree with him or her, think about *why* you disagree and choose an appropriate way to manage your disagreement. Consider, for instance, Paul Graham’s [“hierarchy of disagreement.”](#)

- Strive to think critically. Acknowledge (and at times leverage) your prior knowledge about this topic, but don't be beholden to it.

Coursework will be divided into three categories:

- **In-class writing and discussion** (10% of final grade): During many sessions I will request that you complete a in-class written reflection regarding the topic for the week. This reflection can be about one or more of the readings for the week, or it may address an important public event that relates to material covered in class and/or in your readings. The goal of these short writings is to catalyze in-class discussion and motivate you to think critically about course material. You will also serve as a discussion leader during at least one class, during which time you will share a current events article related to material covered in class and will initiate/facilitate a brief discussion about this article. Finally, we will also perform at least one in-class, hands-on data analysis activity. Note that you do not need any statistics or programming experience prior to taking this course – I will provide you with data and code to get you started
- **Homework** (60% of final grade): Along with course readings, you will be asked to complete several short papers that reflect on the required readings. For these assignments, I do not want to see a summary of what was covered; rather, I would like to see that you are thinking through the issues raised in class and connecting them to topics of interest to you. Perhaps you read an article in the New York Times, saw a TED talk or listened to a podcast that's relevant to something we discussed? Maybe you witnessed an interesting phenomenon on Facebook that made you think back to last Tuesday's discussion? Tell me all about it! Assignments are intended to help you better understand the course material and prepare for your final project. Note that papers should be written in 12 point font and double spaced. I will grade primarily according to the thoughtfulness and clarity of the paper's content.
- **Final project** (30% of final grade): In addition to your homework assignments, you will complete a final paper. For those interested in working with data, you may conduct a short, descriptive analysis of a dataset that may be defined as "big data" and discuss how your data and approach fit within the "big data" paradigm, as well as address the ethical and privacy-related implications of your project. Those who do not wish to analyze data may a.) discuss the challenges and ethical implications of two studies of your choosing that utilize big data, or b.) design a hypothetical project that uses big data to answer social science questions and discuss the methodological limitations and ethical considerations for this project. Final papers should be double spaced in 12 point font. Again, I will grade primarily for thoughtfulness, clarity and thoroughness.

The final percentage to grade point conversion is listed below. Note that assignments turned in late will be given a 20% late penalty for every calendar day following the due date.

% Earned	Final grade	% Earned	Final grade	% Earned	Final grade
>=98%	4.0	85%	3.0	70%	1.5
95%	3.7	80%	2.5	65%	1.0
90%	3.5	75%	2.0	60%	0.7

Academic Dishonesty

I expect students to adhere strictly to the UW's code of conduct. All acts of plagiarism, cheating, or other forms of academic dishonesty will be prosecuted to the full extent allowed under the code. Ignorance of the rules will not be accepted as a valid excuse for breaking them. For more information on the UW Code of Conduct, see <http://www.washington.edu/students/handbook/conduct.html>

Please also review the college's explanation of "Student Academic Responsibility" at <http://depts.washington.edu/grading/pdf/AcademicResponsibility.pdf>.

Academic Accommodations

It is my goal to provide an equally effective learning experience for all students. Please feel free to meet with me any time to discuss how we can best facilitate your academic success. I am happy to provide accommodations as outlined in your letter from Disabled Student Services. To request academic accommodations due to disability, please contact Disabled Student Services, 448 Schmitz Hall, 206-543-8924.

Attendance

Students are expected to attend class regularly. I will not take attendance, but lectures may sometimes include a short, reflective writing assignment. In the case of illness or emergency and *with the permission of the instructor*, in-class reflections can be made up in the form of a one-paragraph, take-home assignment.

Communication:

For simple questions, email is the best way to reach me. I will respond to your messages within 72 hours. For more involved questions or discussions, please use office hours. *Before emailing a question, carefully check that the answer to your question is not already answered in the syllabus, the course website, or the relevant documents.* I reserve the right to not respond to emailed questions that are answered in course documentation

Scheduled Readings and Assignments:

NOTE: To succeed in this course, you must read the assigned sections of the textbook **before** the class session to which it is assigned. I may begin the course with a short written reflection on that day's assigned reading, so in order to receive credit for this reflection and engage in class discussion **please make sure you have read the material before class.**

Week 1: What is big data and what is it not?

- **Tuesday:** Focusing on the data
 - Kitchin, Rob. 2014. "Conceptualizing Data." *The Data Revolution*: 1-26.
- **Thursday:** What do we mean by "big data?"
 - Lohr, Steve. 2012. "The Age of Big Data." *The New York Times*
 - Cukier, Kenneth. 2010. "Data, data everywhere." *The Economist*.

- Kitchin, Rob and Gavin McArdle. 2016. "What Makes Big Data Big Data?: Exploring the ontological characteristics of 26 datasets" *Big Data and Society*.
- **Assignment Due:** N/A

Week 2: Why should sociologists care about big data?

- **Tuesday:** The promises and challenges of using big data
 - Social Science Research Methods: Chapter 5 – Research Design (skim to get you thinking about "traditional" social science research methods and data collection techniques).
 - Goulder, Scott A. and Michael W. Macy. 2014. "Digital Footprints: Opportunities and Challenges for Online Social Research." Annual Review of Sociology.
 - Manovich, Lev. 2011. Trending: the Promises and Challenges of Big Social Data (selection).
- **Thursday:** Research implications of the "Age of Big Data" – does theory matter?
 - Anderson, Chris. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete" *Wired Magazine*.
 - Gonzáles-Bailón, Sandra. "Social Science in the Era of Big Data" *Policy and Internet*.
 - **Assignment due:** Play with Google Trends and Google Correlate and submit a 1-2 page write-up of an interesting trend you discover. Briefly address the limitations and benefits of using Google Trends for this analysis.

Week 3: Addressing data quality: does big data answer your question?

- **Tuesday:** Addressing bias in your data
 - Crawford, Kate. 2013. "Hidden Biases in Big Data." *Harvard Business Review*.
 - Lazar, David et al. 2014 The Parable of Google Flu Trends: traps in big data analysis. *Science*
 - Tufekci, Zeynep. "Big Questions for Social Media Big Data: Representativeness Validity, and Other Methodological Pitfalls."
- **Thursday:** Assessing data quality
 - boyd, danah and Kate Crawford. 2014. "Critical Questions for Big Data." *Information, Communication & Society*.
 - Ang, Chee Siang et al. "Data in the Wild: Some Reflections."
 - **Assignment due:** After reading the abstract of a big data study provided in class, assemble a 1-2 page analysis of the methodological challenges and possibilities of using big data to study this topic.

Week 4: The IRB and human subjects research

- **Tuesday:** Ethics in social science research
 - "The Ethics of Social Research" (pages 60-63).
 - Overview of the "Common Rule" (from The National Institute of Justice)
 - Wikipedia page for the Institutional Review Board.
 - Look over Human Subjects Review Worksheet

- **Thursday:** Ethics in social science research during the “Age of Big Data”
 - Metcalf, Jacob. “Proposed Changes to the Common Rule.” Council for Big Data, Ethics and Society.
 - Zwitter, Andrej. “Big Data Ethics.” *Big Data and Society*.
 - **Assignment due:** Select a case study from <http://bdes.datasociety.net/> and write a 1-2 page summary and discussion of ethical implications of this case study.

Week 5: Are digital traces human subjects/How anonymous is your data?

- **Tuesday:** The challenge of anonymity
 - Wimmer, Andreas and K. Lewis. “Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook.” *American Journal of Sociology*. (pay particular attention to data & methods – skim the rest).
 - Parry, Marc. Harvard Researchers Accused of Breaching Students' Privacy. *Chronicle of Higher Education*.
 - Zimmer, Michael. 2010. “But the Data Are Already Public.” *Ethics in Information Technology*.
 - Ohm, Paul. 2009. “Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization” (pp. 1701-1731 only)
- **Thursday:** Addressing digital traces as human subjects
 - Meyer et al. “Misjudgments will Drive Social Trials Underground.” *Nature News*.
 - Madrigal, Alexis. The Philosopher Whose Fingerprints are All Over the FTC’s New Approach to Privacy.” *The Atlantic*.
 - **Assignment due:** Assemble a 1 to 2 page response regarding ethics of Wimmer and Lewis’s Facebook study

Week 6: Experimental design in the age of big data: The Facebook experiment as a case study

- **Tuesday:** Facebook experiment and what went wrong
 - Kramer, Adam D. et al. 2014. “Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks.” *Science*
 - Rosen, 2014. “Facebook’s Controversial Study is Business as Usual for Tech Companies but Corrosive for Universities”
- **Thursday:** Responses to the Facebook experiment
 - Watts, Duncan J. 2014. “Stop Complaining about the Facebook Study. It's a Golden Age for Research”
 - Crawford, Kate. 2014. “The test we can and should run on Facebook”
 - **Assignment due:** Assemble a 1 to 2 page response to the Facebook study controversy

Week 7: “Everyday Surveillance:” What is collected and what can it tell us?

- **Tuesday:** What is being collected?
 - Lyon, David. “Everyday Surveillance: Personal Data and Social Classifications” *Information, Communication and Society*.
 - Duhigg, Charles. “How Companies Learn Your Secrets. *New York Times*.

- Florida, Richard. "Gentrification through the Eyes of Yelp Reviewers." *The Atlantic*.
- **Thursday:** Who owns this information and how is it used?
 - Goel, Sharad et al. 2010. "Predicting Consumer Behavior with Web Search"
 - Golbeck, Jennifer. 2015. "The Curly Fry Conundrum." TED Talk
 - **Assignment due:** Assemble a 2-3 page write-up that considers one digital trace that you and others leave (for instance, data from your FitBit), how this might be used for social science research, and the ethical implications of using these data.

Week 8: Surveillance and the State

- **Tuesday:** "The Age of Big Data" and what it means for democracy
 - Schneier, Bruce. 2015. *Data and Goliath*. Introduction, Chapters 5 & 6
- **Thursday:** "The Age of Big Data" and what it means for democracy (contd.)
 - Schneier, Bruce. *Data and Goliath*. Chapters 7 & 8
 - **Assignment Due:** N/A (keep working on your final paper!)

Week 9: Wrap up and Final Presentations

- **Tuesday:** Is the "era of big data" a paradigm shift?
 - Kitchin, Rob. "Big Data: New Epistemologies and paradigm shifts." *Big Data and Society*.
- **Thursday:** Final project presentations
 - Prepare a 10 minute presentation of your final paper to discuss with the class
 - **Assignment Due:** Final Paper