

POLITICS OF DATA

ALGORITHMIC CULTURE, BIG DATA, AND INFORMATION WASTE

Syllabus version date: April 5, 2018

Most current syllabus always available [at this link](#) — check version date

“That we are now turning to algorithms to identify what we need to know is as momentous as having relied on credentialed experts, the scientific method, common sense, or the word of God.”

– Tarleton Gillespie (2015)

“Big data” describes everything from weather satellites to Google search histories and Twitter feeds. “Data exhaust” — the “waste” information we leave behind about ourselves whenever we use computers or cellphones — has become a valuable resource for commerce as well as for sociology and political analysis. Algorithms — well-defined procedures for processing information automatically — provide the key to detecting patterns in big data, and turning data exhaust into useful knowledge about people and processes. Taken together, these three phrases capture a major sociotechnical transformation manifested in politically significant phenomena such as “fake news,” Twitter bots, prediction markets, racial profiling, autonomous robotic weapons, Bitcoin, and hacked elections.

These phenomena create unprecedented policy challenges. Effective intervention starts with understanding exactly what these challenges are. The trajectory of this course runs from questions of definition through examples of the often-unexpected, difficult-to-control effects of algorithmic processes and data-centric analysis to the cultural changes associated with algorithmic systems such as Google, Facebook, and Twitter. Finally, we will explore contrasting approaches to the governance of data and algorithmic systems in the United States, the European Union, and the People’s Republic of China.

Readings are drawn from science & technology studies, information science, anthropology, communication, media studies, legal theory, sociology, and computer science, with additional contributions from psychology and philosophy. No particular technical, humanistic, or social scientific background is required, but some familiarity with basic computer science concepts is assumed.

In addition to weekly reading logs, assignments include conducting an “algorithm audit” and writing a policy brief on a topic of your choice.

Prerequisites: none

Requirements and assignments

Summary

- Attendance (you can miss a max of 2 sessions after intro)
- Participation (20 pct)
- Weekly reading responses (15 pct, must be turned in before class, minimum 4 responses required). **Extra credit** for turning in all 6 possible reading responses.
- Serve as lead discussant twice (20 pct)
- Algorithm audit, with a partner (10 pct)
- Policy paper or term paper/project (35 pct)

Details

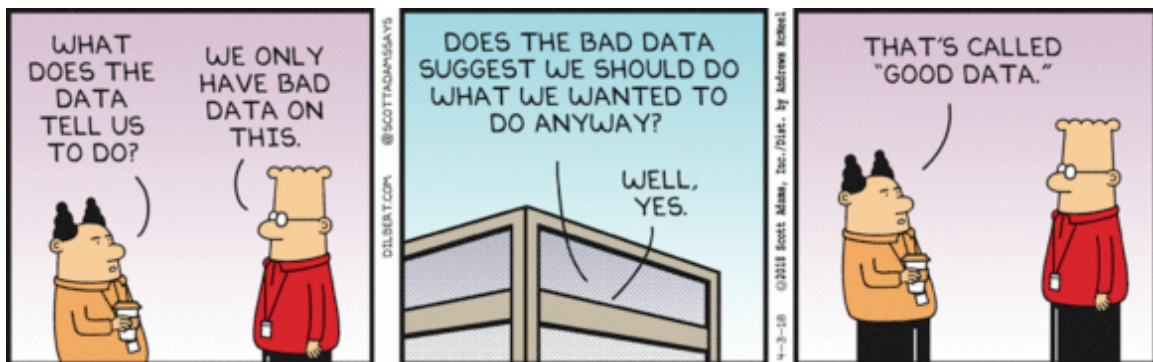
- **Attendance is required.** Discussions are part of the work in this class. Missing more than 2 class sessions during the quarter will negatively affect your final grade, since you can't participate if you're not there. No employer would keep you on if you didn't show up to work 20 percent of the time — and 2 sessions is over 20 percent of this course.
- **Class participation (20 percent of grade).** This is a discussion seminar. Its success depends on the commitment and involvement of *all* participants. **You will be graded on both the regularity and the quality of your participation.**
 - **Cold calls:** I sometimes “cold call” students (ask you direct questions on the readings). I expect your answers to demonstrate your knowledge of the material and your ability to draw interesting connections among the readings. This practice is not intended to single out or embarrass anyone. Its goal is to help you learn to think and talk “on your feet,” a crucial skill required by almost any profession. Please prepare notes on the readings and come to class ready to speak out frequently.
- **Weekly reading responses (15 percent of grade).** A short reading response is due **before** each class session (starting with week 2). See the assignment for details. *These cannot be turned in after the session.* You must complete at least 4 of the 6 possible reading responses in order to pass the course. Extra credit for turning in all 6.
- **Co-lead a class discussion twice (20 percent of grade, 10 pct for each session).** Working with a partner, discussion leaders prepare a discussion plan and a handout for the class (a 1-2 page, *brief* summary of major points from the readings, bullet points only). Maximum time for any presentation is 20 minutes, but other elements (discussion, exercises, small-group work, games) may be added. Discussion leaders meet with the instructor before the class session to go over a draft plan. When you are serving in this role, the weekly reading response is optional.
- **Algorithm audit (10 percent of grade).** With a partner, you'll identify and “audit” an algorithmic process. (See the articles by Sandvig et al. and Diakopolous in the April 18 session.) Since this is potentially a complex project and time is short, in most cases you'll either present an unexecuted audit design, or conduct a “toy” audit. Details to follow.
- **Term research paper or policy paper (35 percent of grade).** A 3000-5000 word paper or policy brief on a data politics issue of your choice. The assignment has five parts.
 1. A 400-500 word prospectus, clearly describing your topic, your approach, and the sources you plan to use, is due April 26.
 2. Meet with the instructor to discuss your idea (April 30 or May 2-3).

3. A **near-complete** first draft (minimum 2500 words) is due May 17.
4. Our final session on May 31 will be a round-robin in which you meet with your peer reviewers to discuss your work. Submit a **complete** draft for peer review by May 29. Peer-review draft papers by two other students. Prepare written notes, or comment in-line on an electronic version. On May 31, discuss your own project and those of your two reviewees.
5. A final version, thoroughly revised and proofread, is due by midnight on June 7. NO LATE PROJECTS.

Required books

Stephens-Davidowitz, Seth. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. HarperLuxe, 2017

Columbia, David. *Politics of Bitcoin: Software as Right-Wing Extremism*. University of Minnesota Press, 2016



Course Schedule

April 5 — Introduction: Learning from Big Data

Stephens-Davidowitz, Seth. *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. HarperLuxe, 2017.

Harford, Tim. "[Big Data: Are We Making a Big Mistake?](#)" *Significance* 11:5 (2014): 14–19

Data & Society Research Institute, 2014 "[Workshop Primer: Algorithmic Accountability](#)"

Sandvig, Christian. "[Seeing the Sort: The Aesthetic and Industrial Defense of 'the Algorithm'](#)," *Media-N: Journal of the New Media Caucus* 10:3 (2014)

Strongly recommended:

Edwards, P.N. "[How to Read a Book](#)"

April 12 — Concepts: algorithmic culture, big data, information waste - Pratyusha and Sidney

Davies, William, "[How Statistics Lost Their Power – And Why We Should Fear What Comes Next](#)," *The Guardian*, 19 January 2017

Smith, Gavin JD. "[Data Doxa: The Affective Consequences of Data Practices](#)." *Big Data & Society* 5, no. 1 (2018): 1–15

Ekbia, Hamid and Bonnie Nardi. "[Heteromation and Its \(Dis\)contents: The Invisible Division of Labor Between Humans and Machines](#)." *First Monday* 19, no. 6 (2014)

Carrigan, Mark, "[Notes on Platform Capitalism](#)" (review of book by Nick Scrnicek), *The Sociological Imagination* (2017)

Strasser, Bruno J and Paul N Edwards. "[Big Data is the Answer... But What is the Question?](#)" *Osiris* 32 (2017): 328–45.

Striphas, T. (2015). "[Algorithmic culture](#)." *European Journal of Cultural Studies*, 18(4-5), 395-412

Recommended:

Ruppert, Evelyn, Engin Isin, and Didier Bigo. "[Data Politics](#)." *Big Data & Society* 4, no. 2 (2017): 1–7.

Plantin, J.-C., C. Lagoze, P. N. Edwards, and C. Sandvig. "[Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook](#)." *New Media & Society* 10 (2016): 1–18.

April 17, 6:15-8:15 PM — Concepts and methods — Nikki and Mikey —

Sandvig, C, K Hamilton, K Karahalios, and C Langbort. "[Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms](#)" (2014)

Diakopoulos, Nicholas. "[Algorithmic Accountability: Journalistic Investigation of Computational Power Structures](#)." *Digital Journalism* 3 (2014): 398–415

Citron, DK and FA Pasquale. "[The Scored Society: Due Process for Automated Predictions](#)." *Washington Law Review* (2014)

Just, Natascha and Michael Latzer. "[Governance By Algorithms: Reality Construction By Algorithmic Selection on the Internet](#)." *Media, Culture & Society* 39, no. 2 (2017): 238–58

Eslami, Motahhare, Karrie Karahalios, Christian Sandvig, Kristen Vaccaro, Aimee Rickman, Kevin Hamilton, and Alex Kirlik. "[First I “like” it, Then I Hide it](#)." CHI Conference (2016): 2371–82

April 26 — Machine learning - Pratyusha and Eric

Video: Zeynep Tufekci, "[We're building a dystopia just to make people click on ads](#)," TED talk (2017), 23 min.

Burrell, J. "[How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms](#)." *Big Data & Society* 3, no. 1 (2016):

Wallach, Hanna. "[Big Data, Machine Learning, and the Social Sciences](#)." *Medium* (2014). Also see the short, related piece by Moritz Hardt, "[How Big Data is Unfair: Understanding Unintended Sources of Unfairness in Data Driven Decision Making](#)." *Medium* (2014).

Domingos, Pedro. "[A Few Useful Things to Know About Machine Learning](#)." *Communications of the ACM* 55, no. 10 (2012): 78-87. This is the most technical article we'll read. If the details bog you down, focus on the discussion.

Mackenzie, A. "[The Production of Prediction: What Does Machine Learning Want?](#)" *European Journal of Cultural Studies* 18, no. 4-5 (2015): 429-45

— *Term project/policy paper: prospectus due* —

May 3 — Algorithm audits (student presentations)

Resources: these are not required, but they may help you in designing your own algorithm audit. I particularly encourage you to watch the video, which Christian Sandvig says is the best example he's seen of an algorithm audit.

Video: Joy Buolamwini and Timimg Gebru, "[Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification](#)," March 2018 (~20 min). There is [an associated paper](#) that may help with details of method.

Datta et al., "[Information Flow Experiments: Determining Information Usage from the Outside](#)" — short summaries of three good audit studies.

[List of algorithm audits](#)

May 10 — Fake news and biased algorithms — Vivien and Mikey

Bessi, Alessandro and Emilio Ferrara. "[Social Bots Distort the 2016 US Presidential Election Online Discussion](#)." *First Monday* 21, no. 11 (2016)

Vlad Barash, Kelly, John, Kollanyi, Bence, Neudert, Lisa-Maria, and Howard, Philip N. "[Polarization, Partisanship and Junk News Consumption Over Social Media in the US](#)," COMPROM Data Memo 2018.1, Project on Computational Propaganda. Oxford Internet Institute, Oxford University (2018)

Wong, Carrie, "[Facebook overhauls News Feed in favor of 'meaningful social interactions'](#)," *The Guardian*, US edition (2018). Explore some of the many links in this short article.

Anderson, Berit, and Brett Horvath, "[The Rise of the Weaponized AI Propaganda Machine](#)," *Medium* (February, 2017)

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort. "[When the Algorithm Itself is a Racist: Diagnosing Ethical Harm in the Basic Components of Software](#)." *International Journal of Communication* 10 (2016)

Moore, Taylor R. "[Trade Secrets and Algorithms as Barriers to Social Justice](#)," Center for Democracy & Technology, Washington DC (2017)

boyd, d., A. E. Marwick, and Karen Levey, "[The Networked Nature of Algorithmic Discrimination](#)," Open Technology Institute (2014)

May 17 — Cryptocurrency and blockchain - Nikki and Eric

Lustig, Caitlin, and Nardi, Bonnie. 2015. "[Algorithmic Authority: The Case of Bitcoin](#)," 48th Hawaii International Conference on System Sciences (HICSS), 743-752

Columbia, David. *Politics of Bitcoin: Software as Right-Wing Extremism*. University of Minnesota Press, 2016. Also read the following reviews of this short book:

- Munger, M., [review of "Politics of Bitcoin,"](#) *The Independent Review: A Journal of Political Economy* 21:4, Spring 2017
- Dupont, Quinn, [review](#) in *Journal of Cultural Economy*
- Optional, longer review, written by a grad student: Osborne, Andrew, "[Chump Change: Decrypting Bitcoin & Blockchain](#)" (2017)

Boucher, Philip. "[How Blockchain Technology Could Change Our Lives: In-Depth Analysis](#)," European Parliament, 2017.

Pisa, Michael and Matt Juden. "[Blockchain and Economic Development: Hype Vs. Reality](#)," *Center for Global Development Policy Paper* 107 (2017)

— **Term project/policy brief: draft due** —

May 24 — Government data and data policy: US, China, EU - Sidney and Vivien

"[M-13-13 — Memorandum for the Heads of Executive Departments and Agencies, Subject: Open Data Policy—Managing Information as an Asset](#)" (2013), memo describing how US open government policies should be interpreted

Tauberer, Joshua, [Open Government Data: The Book](#) (2014). Read at least the following chapters. Also read some of the Case Studies if interested.

- Civic Hacking and Government Data
- A Brief Legal History of Open Government Data
- 14 Principles of Open Government Data
- Example Policy Language

Denis, Jérôme and Samuel Goëta. "Rawification and the Careful Generation of Open Government Data." *Social Studies of Science* 47, no. 5 (2017): 604–29

KPMG International, "[Overview of China's Cybersecurity Law](#)" (February 2017)

Botsman, Rachel, "[Big data meets Big Brother as China moves to rate its citizens](#)," *Wired UK* (21 October, 2017)

European Commission, "[Data Protection in the EU](#)" This website starts with an overview, but it has many links and subsections, including to full-length policy documents. Browse what interests you most.

Wilkinson, MD, et al., "[The FAIR Guiding Principles for Scientific Data Management and Stewardship](#)." *Scientific Data* 3 (2016)

May 31 — Project presentations and final discussion

June 14 — Term project due by midnight. NO LATE PROJECTS. —