

**STS.005/11.155/IDS.057 Data and Society
Spring 2022**

TR 11-12:30 (56-162)



Instructors:

Eden Medina
Program in Science, Technology, and Society
E51-180 | eden@mit.edu | (617) 253-1943
Office Hours: by appointment

Sarah Williams
Department of Urban Studies and
Planning 10-404 | sew@mit.edu |
Office Hours: dusp.mit.edu/oh

TA:

Mariel García-Montes
Program in Science, Technology, and Society
marielgm@mit.edu
Office Hours: by appointment

Description

The course will introduce students to the social, political, and ethical aspects of data science work. It is designed to create reflective practitioners who are able to think critically about how collecting, aggregating, and analyzing data are social processes, and processes that affect people. Students will also be asked to consider how they view their obligation to those who produce the data that data scientists collect and analyze. Students will be expected to participate in class discussions and complete homework assignments. There will be a final project.

Objectives

By the completion of the course students will:

- (1) recognize how social contexts and human values shape the creation of data sets, algorithms, and models;
- (2) recognize how the large-scale collection of data, and the automated systems that utilize these data, can affect people, communities, and organizations.
- (3) identify emergent social, political, legal, and ethical challenges in data science and related fields.
- (4) apply what they have learned in a class project on a current event.

Grading

Final grades will be calculated according to the following:

- 60% Homework assignments
- 30% Final Project
- 10% Participation

Final Project

The final project will encourage students to position the practices and methods of data science in an interdisciplinary context in order to address a highly complex and challenging problem without a clear solution. Students will present their work in a final presentation.

Classroom Behavior

Everyone is expected to conduct themselves professionally during class. This includes arriving on time. Arriving late or using laptops or other devices for entertainment distracts everyone. Professional behavior also entails being respectful of others and their opinions—even and especially when you do not agree with them.

Attendance

Attendance will be taken in class and will form part of the final participation grade.

Attendance is mandatory on presentation days. Students needing to miss a class

presentation because of a serious illness or a personal tragedy will need to receive **advance approval from S³** and provide supporting documentation if requested.

Late Policy

Turning in assignments promptly is important both for keeping up with course material and to keep all students on a relatively level playing field. A late assignment will be accepted up until one week after the original due date for a loss of one letter grade (e.g., an A becomes an A-). After that point, late assignments will receive no credit and will not be accepted.

There will of course be situations where we are willing to bend these rules. We will strive to respect you by taking your requests seriously if you similarly strive to not take advantage of our inclination to respect them. We're all in this weird boat together---let's be good to each other.

Student Conduct, Rights, and Responsibilities

All students are expected to submit their own original work. Plagiarism constitutes using others' ideas, words or images without properly giving credit to those sources. If you turn in any work with your name affixed to it, I assume that work is your own and that all sources are indicated and documented in the text (with quotations and/or citations). For more information on expectations for student conduct please refer to the MIT handbook on academic integrity available online at <http://web.mit.edu/academicintegrity/>.

Student Support Services and Student Disability Services

Every attempt will be made to accommodate qualified students with personal or medical issues that may affect the student's ability to attend class and complete assignments. Students requesting this kind of accommodation must contact and receive verification from Student Support Services (S³). Students may contact S³ by emailing s3-support@mit.edu or calling 617-253-4861. More information is available online at <http://studentlife.mit.edu/s3>.

Every effort will be made to accommodate students with disabilities. Students with disabilities who need accommodations should contact Disability and Access Services (DAS) as soon as possible by emailing das-student@mit.edu or calling 617-253-1674. Students should also let the instructor know of any approved accommodations as soon as possible so that arrangements can be made in a timely fashion. More information is available online at <http://studentlife.mit.edu/das>.

Mental Health

MIT offers a [range of counseling and mental health resources](#) for students. We encourage you to be proactive about taking advantage of them. Do not hesitate to let us know if you're struggling. We are sensitive to the difficulties of our current moment and want to support you as best we can.

Diversity & Inclusion

MIT values an inclusive environment. We hope to foster a sense of community in this classroom and consider this classroom to be a place where you will be treated with respect. We welcome individuals of all backgrounds, beliefs, ethnicities, national origins, gender identities, sexual

orientations, religious and political affiliations – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming, and inclusive environment for every other member of the class. If this standard is not being upheld, please come speak with us immediately.

Writing and Communication Center (WCC)

The Writing and Communication Center offers free one-on-one professional advice from communication experts with advanced degrees and publishing experience. The WCC can help you learn about all types of academic and professional writing and further develop your oral communication skills. You can learn more about the WCC consultations at <http://cmsw.mit.edu/writing-and-communication-center> and register with the online scheduler to make appointments through <https://mit.mywconline.com> - Please note that the WCC hours are offered Monday-Friday, 9:00 a.m.-6:00 p.m., and fill up fast.

Email Etiquette

Please give us at least 24 hours to reply to your emails, and we will do the same for you. Please put the course number in the Subject line and remember to sign your email with your name. We expect the language and structure of your emails to be professional. This includes punctuation, salutations/signature, etc.

Textbooks

Students may wish to acquire the following books. These books may be available via open access or the MIT Libraries.

- ❖ Catherine D'Ignazio and Lauren Klein, *Data Feminism* (Cambridge: MIT Press, 2020).
- ❖ Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences* (London: SAGE Publishers, 2014).
- ❖ Anne Beaulieu and Sabina Leonelli. *Data and Society: A Critical Introduction* (London: SAGE, 2021).
- ❖ Sarah Williams, *Data | Action* (Cambridge: MIT Press, 2020)

These books will appear on the reading list with an (*).

Course Outline

Theme 1: Big Data and Society Overview

February 1 : Introduction to Big Data

- ❖ Class Review

February 3 : Thinking Critically about Big Data

- ❖ Kate Crawford and danah boyd “Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon.” *Information, Communication & Society* 15, no. 5 (2012): 662-679.
- ❖ Chris Anderson, “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete,” *Wired*, June 23, 2008 <https://www.wired.com/2008/06/pb-theory/>
- ❖ Optional: Kate Crawford, Kate, Mary L. Gray, and Kate Miltner, “Critiquing Big Data: Politics, Ethics, Epistemology,” *International Journal of Communication* 8 (2014): 10.

Theme 2 : Building Expert Teams

February 8 : Building Expert Teams : Digital Matatus Building Trust Has Impact

- ❖ Chapter 6: New Data Skills. Beaulieu, Anne, and Sabina Leonelli. *Data and Society: A Critical Introduction**

❖ Assignment #1 Distributed

Case study #1 Stop and Frisk

- ❖ “Case study: New York’s Stop-and-Frisk Policies,” Catalysts for Collaboration, <https://catalystsforcollaboration.org/case-study-floyd-v-city-of-new-york/>

Case Study # 2 : Transport in Nairobi

- ❖ Sarah Williams, Jacqueline Klopp, Peter Waiganjo, Daniel Orwa, and Adam White. "Digital Matatus: Using Mobile Technology to Visualize Informality." *Proceedings ACSA 103rd Annual Meeting: The Expanding Periphery and the Migrating Center* (2015).

Case Study #3 : Automating Public Health

- ❖ David Lazer et al., “The Parable of Google Flu: Traps in Big Data Analysis,” *Science* 343, no. 6176 (March 14, 2014): 1203–5, <https://doi.org/10.1126/science.1248506>.
- ❖ “What We Can Learn From the Epic Failure of Google Flu Trends | WIRED,” October 17, 2018, <https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/>.

February 10 : Historicizing Big Data

- ❖ Chapter 2 : Big Data in Context, Anne Beaulieu and Sabina Leonelli, *Data and Society: A Critical Introduction**
- ❖ Optional: Eden Medina, “Designing Freedom, Regulating a Nation: Socialist Cybernetics in Allende’s Chile,” *Journal of Latin American Studies*.

February 15 : Guest speaker, Jeremy White, graphics editor for *The New York Times*

Theme 3: How Do We Build Data

February 17 : The Ways We Create Data

❖ Assignment #1 DUE

- ❖ Chapter 3 : Characteristics of Data, Beaulieu, Anne, and Sabina Leonelli. *Data and Society: A Critical Introduction**
- ❖ Chapter 1 (only read-up to the Data, Information, Knowledge and Wisdom section.) Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**

Example # 1 : Collection Data from the Bottom

- ❖ Sarah Williams, Chapter 2: Build It! Data is Never Raw, It's Collected, *Data | Action : Modeling for Civic Change**

Example #2 : What Get Collected

- ❖ Catherine D'Ignazio and Lauren Klein, Chapter 2: What Gets Counted, *Data Feminism**

Example 3 # : Cites and Collecting Data Power and Control

- ❖ Sarah Williams, Chapter 1 : Big Data For Cities in Not New, *Data | Action : Modeling for Civic Change**

February 22 : No class; Monday schedule of classes

February 24: Data Creation and Power

- ❖ Introduction, Chapter 1, Catherine D'Ignazio and Lauren Klein, *Data Feminism**
- ❖ David Fidler, ed. *The Snowden Reader* (selections)
- ❖ Joanna Radin, "Digital Natives: How Medical and Indigenous Histories Matter for Big Data," *Osiris* 32 (2017): 43–64.

❖ Assignment #2 Distributed

March 1 : Ways of Thinking about Bias

- ❖ Watch *Coded Bias*, directed by Shalini Kantayya, [Women Make Movies](#), 2020 (documentary). Available on Netflix and online via MIT Libraries.
- ❖ Optional: Helen Nissenbaum and Batya Friedman, "Bias in Computer Systems," ACM Transactions on Information Systems
- ❖ Optional: Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19). Association for Computing Machinery, New York, NY, USA, 59–68.

March 3 : Informed Consent

❖ Assignment #2 Part A due

- ❖ National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*, 1979.

- ❖ Carole Cadwalladr and Emma Graham-Harrison, “Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach,” *Guardian*, March 17, 2018, <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influenc e-us-election>
- ❖ Scott Berinato, “Stop Thinking About Consent: It Isn’t Possible and Isn’t Right,” *Harvard Business Review*, September 24, 2018.
- ❖ Jacob Metcalf and Kate Crawford. "Where are Human Subjects in Big Data Research? The Emerging Ethics Divide." *Big Data & Society* 3, no. 1 (2016).

March 8 : Code of Ethics for Using Data

- ❖ Chapter 9 and 10 : Data Justice and Ethics, Beaulieu, Anne, and Sabina Leonelli. *Data and Society: A Critical Introduction**
- ❖ Matthew Zook et al., “Ten Simple Rules for Responsible Big Data Research,” *PLOS Computational Biology* 13, no. 3 (March 30, 2017): e1005399, <https://doi.org/10.1371/journal.pcbi.1005399>.
- ❖ Sarah Williams, Conclusion : It's How We Work with Data that Really Matters, *Data | Action : Modeling for Civic Change**
- ❖ Catherine D'Ignazio and Lauren F. Klein, “Seven intersectional feminist principles for equitable and actionable COVID-19 data,” *Big Data & Society* 7.2 (2020): 2053951720942544.
- ❖ Greenwood, Faine, et al. "The Signal Code: Ethical Obligations for Humanitarian Information Archives." *Harvard, MA* (2017). Up to page 16 (skim other pages)

March 10 : Presentations of a Curated Dataset

❖ Assignment #2 Part B DUE

Theme 4: How Do We Quantify Data?

March 15 : The Ways We Quantify

- ❖ Chapter 5 : Putting Data to Work, Anne Beaulieu and Sabina Leonelli. *Data and Society: A Critical Introduction**
- ❖ Chapter 1 (only the section called Data, Information, Knowledge, Wisdom), Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**
- ❖ Chapter 6, Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**
- ❖ Chapter 1, Cathy O’Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Broadway Books, 2017).

March 17 : Why We Quantify

- ❖ Chapter 3, Virginia Eubanks, *Automating Inequality*, (New York: St. Martin’s Press, 2018).
- ❖ Martín Tironi & Matías Valderrama (2022) Worth-making in a Datafied World: Urban Cycling, Smart Urbanism, and Technologies of Justification in Santiago de Chile, *The Information Society*, DOI: [10.1080/01972243.2022.2027587](https://doi.org/10.1080/01972243.2022.2027587)

- ❖ Linnet Taylor, “What is Data Justice: The Case for Connecting Digital Rights and Freedoms Globally,” *Big Data & Society*, 2017.

Spring Break March 21th - March 25th

March 29 : Data Visualizations

❖ Assignment #3 Distributed

- Sarah Williams, Chapter 4: Share It!, *Data | Action : Modeling for Civic Change**
- ❖ Jaime Herring et al. “Communicating local climate risks online through an interactive data visualization,” *Environmental Communication* 11.1 (2017): 90-105.
- ❖ Jake Porway, “The Trials and Tribulations of Data Visualization for Good”, Digital Impact, March 15, 2016 <https://digitalimpact.io/the-trials-and-tribulations-of-data-visualization-for-good/>
- ❖ Amar Numanović, “Why policy advocates should pay more attention to data visualization?” Medium, September 12, 2017 <https://medium.com/@numanovicamar/why-policy-advocates-should-pay-more-attention-to-data-visualization-81bba1b8bcea>
- ❖ Yixuan Zhang, Yixuan, et al. "Mapping the Landscape of COVID-19 Crisis Visualizations." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.
- ❖ Crystal Lee: Visualizing Misinformation <https://www.youtube.com/watch?v=vYpGqan2vLw>

March 31: Introduction to the Final Project

- ❖ Linnet Tayloret al. "Data justice and COVID-19." (2020). https://ia801905.us.archive.org/23/items/data-justice-and-covid-19/Data_Justice_and_COVID-19.pdf

Theme 5 :Ground Truthing of Validating Data

April 5 : Data Epistemology

❖ Assignment #3 DUE

- ❖ Chapter 4 : Data, Evidence, and Knowledge, Anne Beaulieu and Sabina Leonelli. *Data and Society: A Critical Introduction**
- ❖ Chapter 3 : Catherine D’Ignazio and Laura Klein, *Data Feminism**
- ❖ Optional: Sabina Leonelli, “What Difference Does Quantity Make? On the Epistemology of Big Data in Biology,” *Big Data and Society*

April 7 : Data Validity

- ❖ Zeynep Tufekci, “Big questions for social media big data: Representativeness, validity and other methodological pitfalls.” *Eighth International AAAI Conference on Weblogs and Social Media*. 2014.
- ❖ Stephen Ansolabehere and Eitan Hersh. "Validation: What big data reveal about survey misreporting and the real electorate." *Political Analysis* 20.4 (2012): 437-459.

April 12 : Guest speaker: Jacqueline Wernimont, Distinguished Chair, Digital Humanities and Social Engagement and Associate Professor of Women's Gender and Sexuality Studies, Dartmouth University

❖ Readings TBD

❖ **Assignment #4 Distributed**

Theme 6: Data Privacy

April 14 : Privacy and Surveillance

- ❖ Chapter 10, Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**
- ❖ Jens Erik-Mai, "Situating Personal Information: Privacy in the Algorithmic Age," in *Human Rights in the Age of Platforms*
- ❖ Shoshanna Zuboff, "It's Not that We've Failed to Rein in Facebook and Google. We've Not Even Tried," *Guardian*, <https://www.theguardian.com/commentisfree/2019/jul/02/facebook-google-data-change-our-behaviour-democracy>

April 19 : Guest Lecture: BU/MIT Technology Law Clinic – U.S. Privacy Law

❖ Robert Solove, "A Brief History of Privacy Law" (selections)

❖ **Assignment #4 DUE**

April 21 : Aggregation and Anonymity

- ❖ Michael Zimmer, "But the Data is Already Public": On the Ethics of Research in Facebook." *Ethics and information technology* 12.4 (2010): 313-325.
- ❖ Paul Ohm, "Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization." *UCLA L. Rev.* 57 (2009): 1701.
- ❖ Gina Kolata, "Your Data Were 'Anonymized'? These Scientists Can Still Identify You," *New York Times* July 23, 2019, <https://www.nytimes.com/2019/07/23/health/data-privacy-protection.html>
- ❖ Thomas Brewster, "120 Million American Households Exposed In 'Massive' ConsumerView Database Leak," *Forbes* December 19, 2017, <https://www.forbes.com/sites/thomasbrewster/2017/12/19/120m-american-households-exposed-in-massive-consumerview-database-leak/#1228bdb37961>
- ❖ Kristin H. Hampton et al. "Mapping health data: improved privacy protection with donut method geomasking." *American Journal of Epidemiology* 172.9 (2010): 1062-1069.

April 26 : Final Project (in class work time)

Theme 7: Share It! : Communicating Data Insights

April 28 : Ethics of Humanitarian Data

- ❖ Kate Crawford and Megan Finn, "The Limits of Crisis Data: Analytical and Ethical Challenges of Using Social and Mobile Data to Understand Disasters," *GeoJournal* 80, no. 4 (August 1, 2015): 491–502, <https://doi.org/10.1007/s10708-014-9597-z>.

- ❖ Linnet Taylor, “The Ethics of Big Data as a Public Good: Which Public? Whose Good?,” *Phil. Trans. R. Soc. A* 374, no. 2083 (2016): 20160126.
<https://royalsocietypublishing.org/doi/full/10.1098/rsta.2016.0126>
- ❖ Sean M. McDonald, “Ebola: A Big Data Disaster-Privacy, Property, and the Law of Disaster Experimentation,” *The Centre for Internet and Society*, no. 2016.01 (2016).
- ❖ Milan, Stefania, and Emiliano Treré. “<? covid19?> The Rise of the Data Poor: The COVID-19 Pandemic Seen From the Margins.” *Social Media+ Society* 6.3 (2020): 2056305120948233.
- ❖ Podcast : Global Data Justice: Episode 7 — Pandemic Politics, Human Rights, and Disinformation <https://globaldatajustice.org/resist-and-reboot/episode-7/>

May 3 : Open Data/Data Transparency

- ❖ Listen to podcast *This American Life*, Episode 584 For Your Reconsideration, “Prologue” and “Knock Knock. Who’s There?: The Truth”
<https://www.thisamericanlife.org/584/for-your-reconsideration>.
- ❖ Benedict Carey and Pam Belluck, “Doubts About Study of Gay Canvassers Rattle the Field,” *New York Times*, May 25, 2015,
<https://www.nytimes.com/2015/05/26/science/maligned-study-on-gay-marriage-is-shakin-g-trust.html>
- ❖ Chapter 3, Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**
- ❖ Chapter 1, B. Goldstein and L. Dyson. *Beyond Transparency: Open Data and the Future of Civic Innovation* (2013).
- ❖ Marijin Janssen, Yannis Charalabidis, and Anneke Zuiderwijk. “Benefits, adoption barriers and myths of open data and open government.” *Information Systems Management* 29.4 (2012): 258-268.

May 5 : Final Project Presentations

May 10 : Final Project Presentations

- ❖ **Final Papers Due May 10 @ 11:59pm**