CS182: Ethics, Public Policy, and Technological Change

Cross-listed as COMM 180, ETHICSOC 182, PHIL 82, POLISCI 182, PUBLPOL 182
Winter 2021 M/W/F 1:30-2:50PM PT
Course Website: cs182.stanford.edu

<table>
<thead>
<tr>
<th>Professor Rob Reich</th>
<th>Professor Mehran Sahami</th>
<th>Professor Jeremy Weinstein</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="mailto:reich@stanford.edu">reich@stanford.edu</a></td>
<td><a href="mailto:sahami@cs.stanford.edu">sahami@cs.stanford.edu</a></td>
<td><a href="mailto:jweinst@stanford.edu">jweinst@stanford.edu</a></td>
</tr>
</tbody>
</table>

**Teaching Team and Course Staff**

We have assembled an interdisciplinary team of Teaching Fellows from across the university, including graduate students with backgrounds in Computer Science, Philosophy, Political Science, Law, and Sociology. Collectively, they bring an array of relevant professional experiences, having served in government, worked at technology companies, and led initiatives to address many of the issues we’ll be discussing. Throughout the course, you will have the opportunity to interact with many different Teaching Fellows.

- **Alena Smith**, Head TA
- **Ece Korkmaz**, Head TA
- **Lorenzo Manuali**, Course Manager
- **Sandra Luksic**, Course Manager
- **Asa Kohrman** (CS)
- **Joy Chen** (Law)
- **Shanduojiao (SJ)**
- **Jiang** (CS)
- **Garrett Walker** (Law)
- **Alex Minsk** (Poli Sci)
- **Ayelet Drazen** (CS)
- **Evan Wisner** (Sym Sys)
- **Sarah Bitter** (Public Policy)
- **Michelle Bao** (CS)
- **Muhammad Yusuf Khattak** (CS)
- **Yilin Wu** (CS)
- **Evani Radiya-Dixit** (Public Policy)
- **Jordan Deasy** (Sustainability Science)
- **David Ludeke** (CS)
- **Julia Kwak** (Phil)
- **Luke Harri Pistol** (Phil)
- **Sonya Kotov** (CS)

You will find a calendar of Office Hours for the course staff on the course website.

If you have questions at any point during the quarter, please feel free to contact any of us. If you are not sure whom to contact, we recommend you start with either one of the Course Managers, Lorenzo Manuali (lmanuali@stanford.edu) and Sandra Luksic (sluksic@stanford.edu), or Head TAs Alena Smith (asvinod@stanford.edu) and Ece Korkmaz (ekorkmaz@stanford.edu).

**Course Description**

Our goal is to explore the ethical and social dimensions of technological innovation. Stanford has a special responsibility to address these topics in light of its role as a seedbed of Silicon Valley. By integrating perspectives from computer science, philosophy, and social science, the course will provide learning experiences that robustly and holistically examine the impact of technology on humans and societies.

The course will challenge students, whatever their choice of major and whatever their career pathway, to think about their role as enablers and shapers of technological change in society. Instead of accepting a common view that what others do with new technologies is their responsibility, students will explore their responsibilities as innovators, designers, coders,
The content of our course will be new. We are building, however, on the long history of ethics and CS at Stanford. The CS department began offering courses in this area in the 1980s, taught by Terry Winograd (Computer Science) and Helen Nissenbaum (Philosophy). Eric Roberts (Computer Science) taught classes on Ethics and Computer Science in the 1990s and 2000s and CS182 continues to be taught regularly in a discussion-focused format.

Targeting students from across multiple disciplines, the course will offer learning opportunities to distinct student populations: giving CS students a greater appreciation of the ethical and policy questions that arise in real-world technical contexts, and students from the humanities and social sciences a deeper understanding of the technical topics underlying many of today’s policy and ethical debates.

Course Topics

The course is structured around five core units, which have been selected primarily for two reasons: (a) to preview critical issues you are likely to play a role in shaping over the next decade and (b) to emphasize topics around which technologists could benefit from greater engagement with domains of knowledge found in other disciplines, including philosophy, social science, and public policy.

<table>
<thead>
<tr>
<th>Unit</th>
<th>Sub-topics of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithmic Decision-Making</td>
<td>- Use of predictive algorithms in public vs. private settings, with emphasis on the criminal justice system</td>
</tr>
<tr>
<td></td>
<td>- Comparative approaches to algorithmic transparency and accountability (e.g., auditing, technological due process)</td>
</tr>
<tr>
<td></td>
<td>- Trade-offs between predictive accuracy and competing values (e.g., fairness, transparency, explainability)</td>
</tr>
<tr>
<td></td>
<td>- Mapping automated systems onto a society characterized by human judgments, informal norms, and formal rules</td>
</tr>
<tr>
<td>The Political Economy of Technology</td>
<td>- Power and ownership over platforms: who exercises power, who has ownership, and why</td>
</tr>
<tr>
<td></td>
<td>- The societal effects of various platforms’ business models</td>
</tr>
<tr>
<td></td>
<td>- Political economies for technologies beyond the standard private market and VC-funding models</td>
</tr>
</tbody>
</table>
| Data Collection, Privacy, and Civil Liberties | Data aggregation, matching, and de-anonymization strategies  
Facial recognition technology (by public and private actors)  
Changing norms and laws around privacy across time and cultures (US, China, Europe), including how people balance privacy vs. other goals  
Consent for different types and uses of data, as well as the idea of companies as “information fiduciaries” |
| --- |
| Artificial Intelligence and Autonomous Systems | Aggregate and distributional consequences of automation (e.g., on labor markets, inequality)  
Role of policy in shaping the impact of AI and automation broadly, as well as specific policy responses (e.g., universal basic income)  
Competitive global landscape of AI development, comparative approaches to governance, and historical precedents for responses (e.g., Asilomar conference) |
| Power of Private Platforms | Transition from an analog to a digital public sphere, with speech and associational rights regulated by companies; virality over veracity in online discourse; tensions between quantity and quality of information; implications for democracy  
Business model concerns, including new conceptions of monopoly and market power of digital platforms, as well as government efforts to promote market competition (e.g., antitrust regulation)  
Technology behind efforts to regulate speech in online communities, including content moderation practices (e.g., banning/deleting speech, upranking/downranking content), frontiers/innovations in speech regulation  
Comparative analysis of how global platforms operate in diverse communities with different speech traditions and politics |
| Blockchain and Decentralized Technical Architectures | A technical explanation of blockchain, its relationship to cryptography, why people claim it creates a ‘public ledger,’ and the importance of such a function  
The relationship between blockchain technologies, the state, and the centralization of power  
Cryptocurrency regulation and how it affects economic efficiency, financial stability, illicit finance, and consumer fraud |

We will also foreground across the entire course two sets of issues that transcend these topics: (1) the political economy of the technology industry and a consideration of alternative corporate and governance structures and (2) the relevance of social categories such as race and gender to the design, development, funding, and deployment of new technologies.
**Pedagogical Structure**

For each of the five modules, we will have a sequence of lectures, discussions, and assignments. Each faculty member will present material relevant to their subject expertise, and an interdisciplinary team of teaching fellows (graduate students from Computer Science, Public Policy, Philosophy, Law School, and elsewhere) will lead weekly small-group sections.

This approach illustrates our strong commitment to a genuine integration of disciplinary approaches. The assignments, course materials, and classroom discussions seek to integrate the technical and non-technical elements of each topic. Throughout the quarter, we will stimulate discussions and provide assignments that require different disciplinary lenses, including coding exercises, policy memos, and philosophical analyses.

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Week 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Promise and Perils</strong></td>
<td><strong>Bringing to Life</strong></td>
</tr>
<tr>
<td><em>intro to topic and competing values at stake</em></td>
<td><em>moderated discussion with experts</em></td>
</tr>
<tr>
<td><strong>Technical Deep Dive</strong></td>
<td><strong>Tensions and Trade-offs</strong></td>
</tr>
<tr>
<td><em>overview of relevant computer science concepts</em></td>
<td><em>interactive discussion on a case study related to the unit (taking place in SECTION)</em></td>
</tr>
<tr>
<td><strong>Rights and Responsibilities</strong></td>
<td><strong>Making Choices</strong></td>
</tr>
<tr>
<td><em>policy implications and social science research</em></td>
<td><em>designing a product/system/policy in light of competing values and trade-offs</em></td>
</tr>
</tbody>
</table>

**Course Requirements, Assignments, and Grading Breakdown**

We have assigned a carefully curated set of readings for each class session. We expect you to complete this reading in advance of each lecture, and to come to section prepared to engage the materials in a facilitated discussion.

Most of the readings for the course are easily accessible through the reading list on the course website. Some readings, however, come from books not available through our university library. **Students must purchase a digital coursepack from the Stanford Bookstore in order to access this copyrighted material.** The bookstore will provide a digital code that allows you to access the coursepack, which contains excerpts from books we’ll be discussing during the quarter. We will provide additional information on how to access the coursepack during the first week of the class.
Assignments and Grading Breakdown

In addition, the course includes five assignments. You will receive more information about each of the assignments well in advance of their due dates.

- Technical assignment – due January 26 at 11:59pm PST
- Philosophy paper – due February 9 at 11:59pm PST [NOTE: WIM students will have an additional revision due February 25 at noon PST]
- Group policy assignment – due March 2 at 11:59pm PST
- Technical assignment – due March 9 at 11:59pm PST
- Final reflection paper – due on March 16 at 11:59pm PST

Grades will be calculated as follows:

<table>
<thead>
<tr>
<th>Non-WIM Students</th>
<th>WIM Students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation – 20%</td>
<td>Participation – 15%</td>
</tr>
<tr>
<td>Technical Assignment 1 – 14%</td>
<td>Technical Assignment 1 – 14%</td>
</tr>
<tr>
<td>Philosophy Paper – 20%</td>
<td>Philosophy Paper (original) – 9%</td>
</tr>
<tr>
<td>Policy Assignment – 20%</td>
<td>Philosophy Paper (revision) – 21%</td>
</tr>
<tr>
<td>Technical Assignment 2 – 6%</td>
<td>Policy Assignment – 20%</td>
</tr>
<tr>
<td>Final Reflection – 20%</td>
<td>Technical Assignment 2 – 6%</td>
</tr>
<tr>
<td></td>
<td>Final Reflection – 15%</td>
</tr>
</tbody>
</table>

Stanford University and its faculty are committed to ensuring that all courses are financially accessible to all students. If you are an undergraduate who needs assistance with the cost of course readings, supplies, materials and/or fees, you are welcome to approach us directly. If you would prefer not to approach us directly, please note that you can ask the Diversity & First-Gen Office for assistance by completing their questionnaire on course textbooks & supplies: [http://tinyurl.com/jpgbarn](http://tinyurl.com/jpgbarn) or by contacting Joseph Brown, the Associate Director of the Diversity and First-Gen Office (jlbrown@stanford.edu). Dr. Brown is available to connect you with resources and support while ensuring your privacy.
Section Sign-ups

In addition to lecture, you must also sign up for a weekly 50 minute section, which will take place every Thursday. You will submit preferences for section times via a form on the CS182 web site (cs182.stanford.edu). **The sign-up form will be available until 5pm, Sunday, January 15th.** After a matching process, your section assignments will be emailed to you. Sections will not be held the first week of classes. Section sign-ups will take place online in the first few days of classes. Note that you should only sign up for sections at the CS182 website (you should not sign-up for sections on Axess). Also note that section sign-ups are not first-come-first-serve; all submissions in by the sign-up deadline will be treated similarly.

Class Participation and Lecture Attendance

Class participation can take a variety of forms, ranging from the obvious (e.g., talking intelligently in class/section) to the less obvious (e.g., sharing articles/podcasts that are relevant to course discussions). **At a minimum, it is crucial that you join class on time, having done the reading, and prepared to talk and engage your fellow classmates.** Because the class will use small group discussions, adequate preparation, willingness to contribute, and capacity for empathetic listening are all required. You are also required to attend a section every week. A portion of your grade will be based on your participation.

Lecture and Section Participation: This course encourages vigorous intellectual exchange, the expression of various viewpoints, and the ability to speak effectively and cogently. Participation includes but is not limited to in-class discussion. As part of the participation grade, the section leaders may assign activities and written assignments.

Except in cases of an OAE accommodation or another valid reason that is brought to the course managers and a student’s course assistant’s attention **before** lecture, **attendance at lectures and sections is mandatory.** If a student has a prolonged illness or a personal situation that might lead to more than one section absence, the student should contact his or her Course Assistant **before** missing section. Under certain conditions, a student may be provided an opportunity to make up work missed in section. In other words, make-up work is at the discretion of the instructor. Note: insufficient section attendance will result in failure of the course.

**In order to be prepared for discussion, it is essential that you come to each lecture having read and understood the materials assigned and having given some thought as to how the readings relate to the course in general. This will allow you to benefit from the class presentations and discussions and in turn prepare yourself to discuss the issues in depth in section. You should come to section with considered views about:**

- what the main claims offered in the texts or case studies are;
- the arguments offered in favor of these claims;
- whether these are good or plausible arguments;
- whether the claim is, all things considered, strong or plausible;
- what alternatives to the claims and arguments exist; and
- whether some alternative is superior to the claim under discussion.

Objections are important. But keep in mind that raising puzzles and problems (even interesting puzzles and problems) for a view is easy: we can be certain in advance that every view will face some problems. Still, we are trying to decide what to think about important issues of enormous
consequence, not playing a game or showing off debater’s skills. The really hard part is to figure out what to think – what we should think – once we understand the range of theoretical options and competing arguments.

Participation in section will be evaluated on the following guidelines, which stress the quality rather than the quantity of contributions.

- **A range:** The student is fully engaged and highly motivated. This student is well prepared, having studied the assigned material, and having thought carefully about the materials’ relation to issues raised in lecture and section. The student’s ideas and questions are substantive (either constructive or critical); they stimulate class discussions. This student listens and responds respectfully to the contributions of other students.
- **B range:** The student participates consistently in discussion. This student comes to section well-prepared and contributes regularly by sharing thoughts and questions that show insight and a familiarity with the material. This student refers to the materials discussed in lecture and shows interest in other students’ contributions.
- **C range:** The student meets the basic requirements of section participation. This student is usually prepared and participates once in a while but not regularly. The student’s contributions relate to the texts and the lectures and offer a few insightful ideas but do not help to build a coherent and productive discussion. (Failure to fulfill satisfactorily any of these criteria will result in a grade of "D" or below.)

**Statement on Open Discourse**

The course is a space for students committed to a rigorous examination of ethics, technology, public policy, and related topics. The course is also a space for respectful, critical inquiry through the free exchange of ideas. Our goal is to come to a greater understanding of – not a consensus on – the issues the course addresses. To that end, this space is defined by mutual respect that allows us, together, to grapple with a range of ideas, evidence, values, and conclusions. The following principles guide our interaction in this space:

- All viewpoints are welcome.
- Treat every member of the course with respect, even if they disagree with another student’s view.
- Treat every claim as open to examination, even if it comes from someone with more experience or expertise than you.
- Reasonable minds can differ on any number of perspectives, opinions, and conclusions.
- Our passions and social and political commitments are welcomed in this space. They are also subject to respectful challenge.
- Some perspectives, opinions, and conclusions are unreasonable or based on falsehoods and should be identified as such.
- No ideas are immune from scrutiny and debate.
- Evidence and reasoning guide our conclusions.

For more information, see:

- Some perspectives, opinions, and conclusions are unreasonable or based on falsehoods and should be identified as such.
- [Free Expression](http://example.com), University of Chicago
- [Class Community Commitments: A Guide for Instructors](http://example.com), CTL
- [Transformative Class Conversations](http://example.com), CTL
The Honor Code

Violating the Honor Code is a serious offense, even when the violation is unintentional. The Honor Code is available at: http://www.stanford.edu/dept/vpsa/judicialaffairs/guiding/honorcode.htm. You are responsible for understanding the University rules regarding academic integrity; you should familiarize yourself with the code if you have not already done so. In brief, conduct prohibited by the Honor Code includes all forms of academic dishonesty, among them copying from another student's work, unpermitted collaboration and representing as one's own work the work of another. If you have any questions about these matters, see your post-doctoral fellow during office hours.

Statement on the Use of Automated Writing/Coding Tools

The development of generative AI (a topic in the class) provides opportunities to support learning but also opportunities to cheat. The use of automated writing or coding tools (e.g., GPT-3 or CoPilot) for submitting assignments is not permitted.

FERPA

Student Record Privacy Policy
http://studentaffairs.stanford.edu/registrar/students/ferpa

Access and Accommodations

Stanford is committed to providing equal educational opportunities for disabled students. Disabled students are a valued and essential part of the Stanford community. We welcome you to our class.

If you experience disability, please register with the Office of Accessible Education (OAE). Professional staff will evaluate your needs, support appropriate and reasonable accommodations, and prepare an Academic Accommodation Letter for faculty. To get started, or to re-initiate services, please visit https://oae.stanford.edu.

If you already have an Academic Accommodation Letter, we invite you to share your letter with us. Academic Accommodation Letters should be shared at the earliest possible opportunity so we may partner with you and OAE to identify any barriers to access and inclusion that might be encountered in your experience of this course.

Additional Resources for Learning

Students who may need an academic accommodation based on the impact of a disability must initiate the request with the Student Disability Resource Center (SDRC) located within the Office of Accessible Education (OAE). SDRC staff will evaluate the request with required documentation, recommend reasonable accommodations and prepare an Accommodation Letter for faculty dated in the current quarter in which the request is being made. Students
should contact the SDRC as soon as possible since timely notice is needed to coordinate accommodations.

**Student Disability Resource Center Office of Accessible Education**

[http://studentaffairs.stanford.edu/oae](http://studentaffairs.stanford.edu/oae)

**The Hume Writing Center** works with Stanford students taking WIM classes and any course that includes writing assignments. In free one-to-one sessions, trained writing consultants help students brainstorm and get started on assignments; learn strategies for revising, editing, and proofreading; and improve organization, flow, and argumentation. We also have digital media consultants who work with students to develop strategies to improve visual and multimodal communication in media such as research posters and PowerPoint and oral communication tutors to help students prepare or refine a presentation. Students can make an appointment with a lecturer or advanced graduate student consultant or drop in to meet with an undergraduate peer tutor. For further information, to see hours and locations, or to schedule an appointment, visit the Hume website at: [http://hume.stanford.edu](http://hume.stanford.edu) / [http://hwc.stanford.edu/](http://hwc.stanford.edu/)

**The Technical Communication Program** (TCP) is a writing and public speaking resource for Stanford students of all levels. TCP is a resource specifically tailored for WIM courses offered in the School of Engineering. Mary McDevitt, Ph.D. (mary.mcdevitt@stanford.edu), Director of TCP, is available to all students for consultation. TCP instructors will also be working with WIM students to provide support on writing assignments. For more information, please visit: [https://engineering.stanford.edu/tcp](https://engineering.stanford.edu/tcp).
Reading List and Course Outline

Below is a schedule of class topics and readings. They are subject to change by the instructors, and all changes will be communicated.

*Required readings that are excerpted from books and are NOT hyperlinked below will appear in the course reader, which you can obtain through the Stanford Bookstore.*

Monday, January 9: Opening session (no pre-reading)

Wednesday, January 11: Computer Science, Optimization, and Ethics

- Existing Codes of Ethics: ACM Code of Ethics and Professional Conduct, Artificial Intelligence at Google – Our Principles, Ethical OS Risk Mitigation Checklist
- Individual profiles: John Luttig, Clarissa Redwine
- “Feynman's Error: On Ethical Thinking and Drifting” by Dan Munro (Dan's blog, November 2018)
- “Optimize What?” by Jimmy Wu

Supplementary:
- Partnership on AI Tenets
- “Of Course Congress Is Clueless About Tech—It Killed Its Tutor” (WIRED, 2016)

Friday, January 13: Algorithmic Decision-Making | Promise & Perils

- “Algorithmic Injustice: A Relational Ethics Approach” by Abeba Birhane. (Patterns, Cell Press. 2021.)

Supplementary:
- The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition by OS Keyes (Proceedings of the ACM on Human-Computer Interaction, Vol. 2, No. CSCW, Article 88. Publication date: November 2018.)
- “Algorithms, Correcting Biases” by Cass Sunstein (Social Research, 2019)
- Video Explainer on Rawls' Original Position (Wireless Philosophy, 2014)

Wednesday, January 18: Algorithmic Decision-Making | Technical Deep Dive

- Handout on “Introduction to Probability and Machine Learning” for Background
- “Machine Bias” by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner (ProPublica, 2016)
- “How we Analyzed the COMPAS Recidivism Algorithm” by Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin (ProPublica, 2016)
- “Can you make AI fairer than a judge? Play our courtroom algorithm game” by Karen Hao and Jonathan Stray (MIT Technology Review, 2019)
● “An Invitation to System-Wide Algorithmic Fairness” (Section 1-3 required; Section 4-5 are supplementary) by Efrén Cruz Cortés and Debashis Ghosh (AIES 2020)

Supplementary:
● “21 Fairness Definitions and Their Politics” by Arvind Narayanan (Tutorial at Conference on Fairness, Accountability, and Transparency, 2018)
● “Algorithmic decision making and the cost of fairness” by Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, Aziz Huq (Proceedings of KDD’17, 2017)
● “It's COMPASlicated: The Messy Relationship between RAI Datasets and Algorithmic Fairness Benchmarks” by Michelle Bao et al. (ArXiv, 2021)
● “Impossibility of What? Formal and Substantive Equality in Algorithmic Fairness” by Ben Green (SSRN, 2021)

Thursday, January 19 (SECTION): Introductory Section

Friday, January 20: Algorithmic Decision-Making | Rights & Responsibilities
● “If You Give a Judge a Risk Score: Evidence from Kentucky Bail Decisions” by Alex Albright
● Alan Gerber and Donald Green, Field Experiments, Chapter 1

Supplementary:
● “Improving Refugee Integration through Data-Driven Algorithmic Assignment” by Kirk Bansak, Jeremy Ferwerda, Jens Hainmueller, Andrea Dillon, Dominik Hangartner, Duncan Lawrence, Jeremy Weinstein (Science, 2018)
● “From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation” by Scott E. Carrell, Bruce I. Sacerdote, and James E. West (Econometrica, 2013)
● “Randomized Controlled Field Trials of Predictive Policing” by Mohler et al, (Journal of the American Statistical Association, 2015)

● “Discrimination in the Age of Algorithms” by Kleinberg et al, pp. 1-6 (NBER, 2019)
● “Algorithmic Impact Assessments: Toward Accountable Automation in Public Agencies” by Dillon Reisman, Jason Schultz, Kate Crawford, Meredith Whittaker, pp. 7-20 (AI Now Institute, 2018)
Discriminating Systems: Gender, Race, and Power in AI by Sarah West, Meredith Whittaker, and Kate Crawford (AI Now Institute, 2019)

Supplementary:
- Frank Pasquale, The Black Box Society, all of ch. 5-6
- “Data and Society Report Algorithmic Impact Assessment” (Policy Brief), by Emmanuel Moss, Elizabeth Anne Watkind, Ranjit Singh, Madeleine Clare Elish, and Jacob Metcalf (2021)

Wednesday, January 25: Political Economy Module | Optimization and History
Note: Technical Assignment #1 is due tomorrow at 11:59 pm PST!
- Forgotten Women Who Built the Internet by Mark Richers.
- Chapter 5, “The State Behind the iPhone” from The Entrepreneurial State by Mariana Mazzucato.

Supplementary

Thursday, January 26 (SECTION): Algorithmic Decision-Making | Bringing to Life
- Case Study: Algorithmic Decision-Making and Accountability

Supplementary:
- When an Algorithm Helps Send You to Prison by Ellora Israni (New York Times, 2017)

Friday, January 27: Political Economy Module | Power and Ownership
- Chapter 4, "Nothing For Us Without Us" from Design Justice by Sasha Costanza-Shock (only read the first section - starting at “Introduction” and ending with Design Process: From Participation to Accountability to Ownership. (2020)
- Interview with Veena Dubal on Prop 22 as a “New Racial Wage Code” (2022)
- “The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power” by Pawel Popie. (Communication, Culture, and Critique, 2018)

Supplementary
● Scroll through techinquiry.org and look up your choice of nonprofits, universities, and company to see the revolving door between technology and the U.S. military.

● "The Steep Cost of Capture" by Meredith Whittaker. 2021. Interactions. (Pages: 6)

● "Digital colonialism: US empire and the New Imperialism in the Global South" by Michael Kwet


● "Combatting Anti-Blackness in the AI Community" by Devin Guillery (arXiv, 2020)

Monday, January 30: Data Collection, Privacy, Civil Liberties | Promise & Perils

● Nothing to Hide, pp. 21-32, by Daniel Solove (Yale University Press, 2013)


● Discipline and Punish, ch. 3 "Panopticism" by Michel Foucault

● Stanford Administrative Guide, 6.1.1 Privacy and Access to Electronic Information

● "How Surveillance has Always Enforced Racism" by Sidney Fussell (WIRED, 2020)

● "How Capitalism—Not a Few Bad Actors—Destroyed the Internet" by Matthew Crain. (Boston Review, 2022)

Supplementary:

● “A Contextual Approach to Privacy Online” by Helen Nissenbaum (Daedalus 140 (4), Fall 2011)

● “Limitless Worker Surveillance” by Ifeoma Ajunwa, Kate Crawford, and Jason Schultz (California Law Review, 2016)

● "Privacy Harms" by Danielle Keats Citron & Daniel J. Solove (Forthcoming, Boston University Law Review, 2022)

● "Facial Recognition: What Happens When We’re Tracked Everywhere We Go?" by Kashmir Hill (New York Times, 2021)

● "The Constant Boss: Work Under Digital Surveillance" by Aiha Nguyen (Data and Society, 2021)

Wednesday, February 1: Data Collection, Privacy, Civil Liberties | Technical Deep Dive

● “Private traits and attributes are predictable from digital records of human behavior” by Michal Kosinski, David Stillwell, and Thore Graepel (PNAS, 2013)

● "Differential Privacy: A Primer for a Non-technical Audience" by Alexandra Wood et al. (Vanderbilt Journal of Entertainment & Technology Law, 2018), pp. 211-214 (Executive Summary) and pp. 225-246 (Sections III and IV)

● "What is Differential Privacy?" by Matthew Green (A Few Thoughts on Cryptographic Engineering, 2016)

● "Harvard Researchers Identify Accuracy Concerns in Census Bureau’s New Privacy System" by Kate N. Guerin (The Harvard Crimson, 2021)

● "Why ‘Anonymous’ Data Sometimes Isn’t" by Bruce Schneier (WIRED, December 2007)

Supplementary:

● “The Promise of Differential Privacy: A Tutorial on Algorithmic Techniques” by Cynthia Dwork (Microsoft Research, 2011)

Thursday, February 2 (SECTION): Data Collection, Privacy, Civil Liberties | Tensions & Trade-offs

● Case Study: Facial Recognition

● Individual profiles: Courtney Bowman

● “Civil Society Letter to Amazon on Facial Recognition” (Human Rights Watch, 2019)
Supplementary

- “The End of Trust” from McSweeney’s and Electronic Frontier Foundation
- “The Perpetual Line-Up: Unregulated Police Face Recognition In America” from Center on Privacy and Technology at Georgetown Law
- Podesta report “Big Data: Seizing Opportunities, Preserving Values” (especially pp. 58-68)
- “Report on the Telephone Records Program Conducted under Section 215” (Privacy And Civil Liberties Oversight Board, 2014)
- “Facial recognition technology: The need for public regulation and corporate responsibility” by Brad Smith (Microsoft, 2018)
- “Most US government agencies are using facial recognition” by Russell Brandom (The Verge, 2021)

Friday, February 3: Data Collection, Privacy, Civil Liberties | Rights & Responsibilities

- “Privacy and human behavior in the age of information” by Alessandro Acquisti, Laura Brandimarte, and George Loewenstein (Science, 2015)
- “Privacy and Information Sharing” by Lee Rainie and Maeve Duggan, pp. 1-8 (skim the rest), (Pew Research Center, 2016)
- “Americans feel the tensions between privacy and security concerns” by Shiva Maniam (Pew Research Center, 2016)

Supplementary:

- “Privacy and Data Protection in an International Perspective” by Lee A. Bygrave, sections 3-5 (Scandinavian Studies in Law, 2010)

Monday, February 6: Data Collection, Privacy, Civil Liberties | Bringing to Life

We will be joined in class by Meredith Whittaker (President of Signal) and TBD

- “Why Signal won’t compromise on encryption, with president Meredith Whittaker” by Nilay Patel, (2022), The Verge.

Supplementary

- “Meredith Whittaker’s Signal And A Progressive Vision For Tech” (2022), Hosted by Alexis Madrigal, KQED Forum.

Wednesday, February 8: Data Collection, Privacy, Civil Liberties | Making Choices

Note: Your philosophy paper is due tomorrow at 11:59pm PST!

- GDPR, Art. 5 “Principles relating to processing of personal data
• “A Design for Public Trustee and Privacy Protection Regulation” by Priscilla M. Regan (Seton Hall Legislative Journal, 2020)

Supplementary:
• “Jaron Lanier Fixes the Internet” by Jaron Lanier and Adam Westbrook (video series) (New York Times, 2019)
• “Information Fiduciaries and the First Amendment” by Jack Balkin (UC Davis Law Review, 2016)
• "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI" by Sandra Wachter and Brent Mittelstadt, pp. 1-18, 78-85 (Columbia Business Law Review, Forthcoming)
• “We May Own Our Data, but Facebook Has a Duty to Protect It” by Nathan Heller (New Yorker, 2018)
• “State Privacy Legislation Stalls Despite High Hopes” by Ashley Gold (The Information, 2019)
• “Privacy, Autonomy, and the Dissolution of Markets” by Kiel Brennan-Marquez and Daniel Susser (Knight Institute, 2022)

Friday, February 10: AI and Autonomous Systems | Promise & Perils
• “The Experience Machine” by Robert Nozick (Excerpt from Anarchy, State, and Utopia, 1974)
• 2022 AI Index Report by Stanford HAI. (Focus on chapter 4, skim the rest.)
• “Nick Bostrom – A Frightened Optimist On The Future of Humanity” by Nick Bostrom and Paula Erizanu (Institute of Art and Ideas, 2019)

Supplementary:
• “Oil is the New Data” by Zero Cool (Logic Magazine, 2019)
• “The Vulnerable World Hypothesis” by Nick Bostrom (2018)
• “Artificial Intelligence, Humanistic Ethics” by John Tasioulas (2021)

Monday, February 13: AI and Autonomous Systems | Technical Deep Dive
• “Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report.” (Stanford, 2021): Read About AI100 (pp. 4), SQ6 (pp. 33-37), SQ7 (pp. 37-43), SQ8 (pp. 43-47), SQ9 (pp. 48-52), SQ10 (pp. 53-56), SQ11 (pp. 56-60), WQ1 (pp. 63-67), Conclusions (pp. 71)
• “Society-in-the-Loop” by Iyad Rahwan (ArXiv, 2017)
• “Introduction” and “Automation 2.0” from Smart Machines and Service Work: Automation in the Age of Stagnation by Jason Smith. 2020, pgs. 7-15 (Pages: 8).

Supplementary:
• “Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report.” (Stanford, 2021): Remainder of report not specified in the required reading above.
• Artificial Intelligence in Society, Chapter 4 [Public Policy Considerations], (OECD, 2019)
Wednesday, February 15: AI and Autonomous Systems | Rights & Responsibilities

- “Artificial Intelligence, Automation and Work” by Daron Acemoglu and Pascual Restrepo, pp. 1-15 (NBER, 2018)
- Chapter 2: “Labor” from *The Atlas of AI* by Kate Crawford, particularly pgs. 70-91, 104-110. 2021

Supplementary:
- “Liability and Regulation of Autonomous Vehicle Technologies” by Nidhi Kalra, James M. Anderson and Martin Wachs, pp. 17-36 (RAND, 2009)
- Andrew McAfee & Erik Brynjolfsson, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*
- “The American Corporation is in Crisis—Let's Rethink It” by Lenore Palladino (Boston Review, 2019)
- “New World Order” by Erik Brynjolfsson, Andrew McAfee, and Michael Spence (Foreign Affairs, 2014)

Thursday, February 16 (SECTION): AI and Autonomous Systems | Tensions & Trade-offs

- Case Study: Autonomous Vehicles
- “The Pursuit of the Ideal” by Isaiah Berlin (ch. 1 from *The Crooked Timber of Humanity*)

Supplementary:
- “Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations” by Johannes Himmelreich (Ethical Theory and Moral Practice, 2018)
- ”Isaiah Berlin: Against Dogma” by Henry Hardy (Times Literary Supplement, 2020)
- “Pluralism and Incommensurable Goods” by Elizabeth Anderson (in *Value in Ethics and Economics*)

Friday, February 17: AI and Autonomous Systems | Bringing to Life

*We will be joined in class by Dr. Margaret Mitchell (Hugging Face) and Dr. Pattie Maes (MIT Media Lab)*

Supplementary

Wednesday, February 22: AI and Autonomous Systems | Making Choices

- “Taxes, Automation and the Future of Labor” by Daron Acemoglu, Andrea Manera, Pascual Restrepo (MIT, 2020)
Supplementary:
- “The Short-Term Impact of Unconditional Cash Transfers to the Poor” by Johannes Haushofer and Jeremy Shapiro (The Quarterly Journal of Economics, 2016)
- “Who Really Stands to Win from Universal Basic Income” by Nathan Heller (New Yorker, 2018)
- Y Combinator Basic Income Study (browse website)
- “A Layered Model for AI Governance” by Urs Gasser and Virgilio A.F. Almeida (IEEE Internet Computing, 2017)
- “Is effective regulation of AI possible? Eight potential regulatory problems” by John Danaher (Philosophical Disquisitions, 2014)
- “Perspectives on Issues in AI Governance,” Google
- AI Superpowers, Chapter 4, by Kai-Fu Lee (Houghton Mifflin Harcourt, 2018)

Friday, February 24: Power of Private Platforms | Promise & Perils
Note [WIM students only]: Your revised philosophy paper is due tomorrow (the 25th) at noon PST!
- Chapter 2 of On Liberty by John Stuart Mill (republished by Heterodox Academy, 2019)
- “A Declaration of Independence of Cyberspace” by John Perry Barlow (1996)
- “Democracy’s Dilemma” by Henry Farrell and Bruce Schneier (Boston Review, 2019)

Supplementary:
- “It’s the (Democracy-Poisoning); Golden Age of Free Speech” by Zeynep Tufekci (WIRED, 2018)
- “Is the First Amendment Obsolete?” by Tim Wu, pp. 2-17 (Columbia Public Law Research Paper, 2018)
- Tim Wu, The Curse of Bigness, Intro + ch. 7
- Stanford Encyclopedia of Philosophy, Freedom of Speech

Monday, February 27: Power of Private Platforms | Technical Deep Dive
- “Exposure to ideologically diverse news and opinion on Facebook” by Eytan Bakshy, Solomon Messing, Lada A. Adamic (Science, 2015)
- “How the biggest decentralized social network is dealing with its Nazi problem” by Adi Robertson (2019). The Verge.

Supplementary:
- “Personalized News Recommendation Based on Click Behavior” by Jiahui Liu, Peter Dolan, and Elin Ranby Pedersen (Google, 2009)
- “Filter Bubbles, Echo Chamber, and Online News Consumption” by Seth Flaxman, Sharad Goel, and Justin M. Rao (Public Opinion Quarterly, 2016)
- “I Was the Head of Trust and Safety at Twitter. This Is What Could Become of It.” by Yoel Roth (New York Times, 2022)
Wednesday, March 1: Power of Private Platforms | Rights & Responsibilities
Note: Your group policy memo is due tomorrow at 11:59pm PST!
- “Democracy and the Digital Public Sphere” by Josh Cohen and Archon Fung from *Digital Technology and Democratic Theory*, Bernholz, Landemore, Reich, eds (University of Chicago Press, 2021)

Supplementary:
- “Trends in the Diffusion of Misinformation on Social Media” by Hunt Allcott, Matthew Gentzkow, Chuan Zu (2018)
- “Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature” by Josh Tucker (Hewlett Foundation, 2018)
- “Free Speech is a Triangle” by Jack Balkin (Columbia Law Review, 2018)
- “Tech Platforms and the Knowledge Problem” by Frank Pasquale (American Affairs, 2018)

Thursday, March 2 (SECTION): Power of Private Platforms | Tensions & Trade-offs
- **Case Study: Platforms**
- Individual profile: Minnie Ingersoll

Friday, March 3: Power of Private Platforms | Making Choices
- “Dealing with Disinformation: Evaluating the Case for Amendment of Section 230 of the Communications Decency Act” by Tim Hwang, pp. 252-285 in *Social Media and Democracy* Edited by Nathaniel Persily and Joshua A. Tucker (Cambridge University Press, 2020)
- “Free Speech Is Not the Same As Free Reach” by Renee DiResta (WIRED, 2018)
- “From Private Bads to Public Goods: Adapting Public Utility Regulation for Informational Infrastructure” by K. Sabeel Rahman and Zephyr Teachout (Knight First Amendment Institute, 2020)

Supplementary:

Monday, March 6: Blockchain and Decentralized Technology | Technical Deep Dive
- “But how does bitcoin actually work?” by 3Brown1Blue (2017)
• “Understanding Zero-Knowledge Proofs Through Illustrated Examples” by 0xSage (2019).

Supplementary:

Wednesday, March 8: Blockchain and Decentralized Technology | Promises & Perils
Note: Technical Assignment #2 is due tomorrow at 11:59pm PST!

• Leviathan, §13 and §17 by Thomas Hobbes (1651). The selection contains more — the other sections are supplementary.
• The Network State, §2.4 by Balaji Srinivasan

Supplementary:
• “Blockchain as a confidence machine: The problem of trust & challenges of governance,” by Primavera De Filippi, Morshed Mannanc, and Wessel Reijersd (2020).
• Leviathan, §14-15, 16, 18-19 by Thomas Hobbes (1651).
• The Network State, §2.2 by Balaji Srinivasan
• “Effective Altruism Is Pushing a Dangerous Brand of ‘AI Safety’” by Timnit Gebru (2022). WIRED.

Friday, March 10: Blockchain and Decentralized Technology | Rights & Responsibilities

• “Written Testimony; Shane T. Stansbury; Duke University School of Law; United States Senate Committee on Banking, Housing, and Urban Affairs; Hearing on “Understanding the Role of Digital Assets in Illicit Finance” by Shane Stansbury (2022).
Supplementary

- *The Denationalization of Money* by Friedrich Hayek (1976). Especially sections 1, 3-4, 8-10, 12-13, 16, 18-19, 21, 24-25.
- "Reports show scammers cashing in on crypto craze" by Emma Fletcher (FTC, Data Spotlight) (2022).

Monday, March 13: Blockchain and Decentralized Technology | Bringing to Life
We will be joined in class by TuongVy Le and Dave Morin.
- Twitter thread, @TuongvyLe12

Wednesday, March 15: Political Economy Module | Class 3: Alternative Political Economies
Note: Your Final Reflection Paper is due tomorrow at 11:59pm PST!
- "The Case for Digital Public Infrastructure" by Ethan Zuckerman. 2020. Knight First Amendment Institute at Columbia University
- "From Private Bads to Public Goods: Adapting Public Utility Regulation for Informational Infrastructure" by Sabeel Rahman and Zephyr Teachout. February 4, 2020,
- "Let Users Own the Tech Companies They Help Build" by Morshed Mannan and Nathan Schneider. Wired. 2021.
- "Nationalize Amazon" by Paris Marx. (Jacobin, 2020).

Supplementary
- "Exit to Community: Strategies for Multi-Stakeholder Ownership In the Platform Economy" by Morshed Mannan and Nathan Schneider. 2021. Georgetown Law Tech Review. Pages 1-21, skim the rest.

Friday, March 17: Wrapping Up
- "The Ones Who Walk Away from Omelas" by Ursula K. Le Guin (1973)
- "With Great Tech Comes Great Responsibility" by Mozilla Foundation, pp. 9-10 (2020)
- "Confucius and the Whistleblower" by Peter Wei, Palladium Magazine

Supplementary
- "Uncanny Valley" by Anna Wiener (n+1, 2016)
- "What is Technology?" by Saffron Huang (Letters to a Young Technologist)
- "Value By Instrumentalization" by Jasmine Wang (Letters to a Young Technologist)
- "Study the Past, Create the Future" by Matthew Jordan (Letters to a Young Technologist)
- "To be a Technologist is to be Human" by Saffron Huang, Maran Nelson (Letters to a Young Technologist)
● “It’s Time to Govern” by Anna Mitchell (Letters to a Young Technologist)