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 1:30-3PM PT
Course Website: cs182.stanford.edu

Professor Rob Reich
reich@stanford.edu
Professor Mehran Sahami
sahami@cs.stanford.edu
Professor Jeremy Weinstein
jweinst@stanford.edu

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.

Chloe Stowell, Head TA
Lorenzo Manuali, Course Manager
Sandra Luksic, Course Manager
Daniel Slate
Alessandro Vecchiato
Elena Berman
Asa Koehrman
Ece Korkmaz
Adrian Liu
Crystal Liu
Jeffrey Propp
Chuqing (Cathy) Yang
Amber Yang
Shreya Venkat
Yilin Wu
Shanduo Jiao Jiang

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 TA Chloe Stowell (stowell@stanford.edu).

Course Description

Our goal is to explore the ethical and social impacts 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, engineers, corporate leaders, policymakers, citizens, and consumers. With every new innovation, students will ask: What am I enabling others to do? What responsibilities does this imply for me as an innovator, a citizen, and a human being?

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 of several hundred students on Ethics and Computer Science in the 1990s and 2000s and CS182 continues to be taught regularly in a smaller, 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 four 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>
</table>
| **Algorithmic Decision-Making** | • Use of predictive algorithms in public vs. private settings, with emphasis on the criminal justice system  
  • Comparative approaches to algorithmic transparency and accountability (e.g., auditing, technological due process)  
  • Trade-offs between predictive accuracy and competing values (e.g., fairness, transparency, explainability)  
  • Mapping automated systems onto a society characterized by human judgments, informal norms, and formal rules |
| 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)  
● Autonomous weapons, including discussion of AI for military use, and autonomous vehicles  
● 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 |
| Technology, Inclusion, and Inequality | ● Issues of diversity and culture within the tech industry—who is represented at these companies and for whom are they creating products and services  
● Who is included in the prosperity created by tech companies and who is not—i.e. gentrification, gig workers  
● How tech and its specific business models are prone to market concentration and exacerbate U.S. economic inequality  
● Tech worker-led organizing to create change |

We will also foreground across the entire course a set of important issues that transcend these topics, including corporate business models (e.g., attention economy) and practices (e.g., hiring)
in many technology companies. The political economy of the tech industry and concerns about hiring practices are too important to relegate to a single session. For example, with respect to issues of diversity and inclusion in the technology industry, we will address a number of related but independent elements: concerns about justice that hiring practices and talent pipelines raise (e.g., the fair treatment of women and minorities), the epistemic advantages of diversity when it comes to identifying problems worth solving and then creating solutions, and the ways that organizational choices and policies shape individual and collective experiences.

**Pedagogical Structure**

For each of the five units, 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>&lt;br&gt;intro to topic and competing values at stake</td>
<td><strong>Bringing to Life</strong>&lt;br&gt;moderated discussion with experts</td>
</tr>
<tr>
<td><strong>Technical Deep Dive</strong>&lt;br&gt;overview of relevant computer science concepts</td>
<td><strong>Tensions and Trade-offs</strong>&lt;br&gt;interactive discussion on a case study related to the unit (taking place in SECTION)</td>
</tr>
<tr>
<td><strong>Rights and Responsibilities</strong>&lt;br&gt;policy implications and social science research</td>
<td><strong>Making Choices</strong>&lt;br&gt;designing a product/system/policy in light of competing values and trade-offs</td>
</tr>
</tbody>
</table>

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

We have assigned a modest and 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 on algorithmic decision-making – due January 21
- Group policy assignment on privacy – due February 2
- Philosophy paper on autonomous systems – due February 16 [NOTE: WIM students will have an additional revision due later]
- Technical assignment on platforms and social networks – due February 27
- Final reflection paper — due on March 11

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 (including revision) – 30%</td>
</tr>
<tr>
<td>• Policy Assignment – 20%</td>
<td>• Policy Assignment – 20%</td>
</tr>
<tr>
<td>• Technical Assignment 2 – 6%</td>
<td>• Technical Assignment 2 – 6%</td>
</tr>
<tr>
<td>• Final Reflection – 20%</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 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 on Sunday, January 9th. 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 case studies, 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.

As of now, the first two weeks of this class (including all lectures and sections) will be completely virtual. The virtual/in-person nature of the class is subject to change pending any new announcements from Stanford administration.

After the first two weeks, the following policy applies. Students with appropriate OAE accommodations or needing to quarantine during the quarter will be able to access lecture recordings upon request to the course managers. Should a student require accessibility accommodations, please contact the course managers indicating so, and lecture recordings can be made available on an ongoing basis. Otherwise, in case of illness or another suitable reason, specific lecture recordings will be made available upon request to the course managers. Please email the course managers with any questions.

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 the 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.)

(NEW) **Partially Asynchronous Course Policy**
Students may, for a medical or other suitable reason, request to take the rest of the course partially asynchronously. To do so, send an email to the course managers and Head TA specifying a reason for an exception to the in-person, synchronous course attendance policy. In the (partially) asynchronous version of the class, students will have access to recordings of the
lectures. In order to earn credit for participation, starting in Week 3, students will have to submit written responses to the lecture and readings, with a length and frequency to be determined in consultation with the course staff. **Asynchronous students must still attend all sections (they are on Thursdays) synchronously in order to pass the class**, barring a suitable reason (e.g., illness).

**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**, University of Chicago
- **Class Community Commitments: A Guide for Instructors**, CTL
- **Transformative Class Conversations**, CTL
- **Microaggressions and Microaffirmations Series**, UC Davis Center for Educational Excellence.
- **Protected Identity Harm Reporting**, Dean of Students
- **Implicit Bias and Microaggressions**, 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](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.

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

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

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://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.
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 3: Opening session (no pre-reading)

Wednesday, January 5: Codes of Ethics, Disciplinary Perspectives, and Case Studies

- Existing Codes of Ethics: [ACM Code of Ethics and Professional Conduct](https://www.acm.org/about/ethics), [Artificial Intelligence at Google – Our Principles, Ethical OS Risk Mitigation Checklist](https://ai.google/about/principles/)
- Individual profiles: Ellora Israni, John Luttig, Henry Tsai, Chris Cox
- “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](https://partnershiponai.org)
- “Of Course Congress Is Clueless About Tech—It Killed Its Tutor” (WIRED, 2016)
- “Heidegger on Gaining a Free Relation to Technology” by Hubert Dreyfus (1997)

Friday, January 7: Algorithmic Decision-Making | Promise & Perils


Supplementary:
- “Algorithms, Correcting Biases” by Cass Sunstein (Social Research, 2019)
- Video Explainer on Rawls’ Original Position (Wireless Philosophy, 2014)

Monday, January 10: 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)

Supplementary:
- “21 Fairness Definitions and Their Politics” by Arvind Narayanan (Tutorial at Conference on Fairness, Accountability, and Transparency, 2018)
• “Fair and Unbiased Algorithmic Decision Making: Current State and Future Challenges” by Songul Tolan (ArXiv, 2019)
• “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)

Wednesday, January 12: 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)

Thursday, January 13 (SECTION): Introductory Section

Friday, January 14: Algorithmic Decision-Making | Making Choices
• Frank Pasquale, The Black Box Society, pp. 140-143, 147-153 [jump to “The Lawful Use of Data” on p. 147] (Harvard University Press, 2016)
• “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)

Supplementary:
• Frank Pasquale, The Black Box Society, all of ch. 5-6
Wednesday, January 19: Algorithmic Decision-Making | Tensions & Trade-offs  
*We will be joined in class by Joaquin Candela and Marietje Schaake*

Thursday, January 20 (SECTION): Algorithmic Decision-Making | Bringing to Life  
- Case Study: Algorithmic Decision-Making and Accountability

Supplementary:  

Friday, January 21: Data Collection, Privacy, Civil Liberties | Promise & Perils  
*Note: Technical Assignment #1 is due today!*
- *Nothing to Hide*, pp. 21-32, by Daniel Solove (Yale University Press, 2013)
- *Why Democracy Needs Privacy,* by Carissa Veliz (Boston Review, 2021)
- *Discipline and Punish*, ch. 3 *Panopticism* by Michel Foucault
- *Stanford Administrative Guide, 6.1.1 Privacy and Access to Electronic Information*

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)
- *How Surveillance has Always Enforced Racism* by Sidney Fussell (WIRED, 2020)
- *Privacy Harms* by Danielle Keats Citron & Daniel J. Solove (Forthcoming, Boston University Law Review, 2022)

Monday, January 24: 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)
- *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)
Wednesday, January 26: Data Collection, Privacy, Civil Liberties | Bringing to Life
We will be joined in class by Sridhar Ramaswamy

Thursday, January 27 (SECTION): Data Collection, Privacy, Civil Liberties | Tensions & Trade-offs
- Case Study: Facial Recognition
- Individual profiles: Andrew Grotto, 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, January 28: Data Collection, Privacy, Civil Liberties | Rights & Responsibilities
- Shoshana Zuboff, The Age of Surveillance Capitalism, Chapter 18 (PublicAffairs, 2019)
- "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)
- "Nudging Privacy: The Behavioral Economics of Personal Information" by Alessandro Acquisti (IEEE Security & Privacy, 2009)

Monday, January 31: Data Collection, Privacy, Civil Liberties | Making Choices
- 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)
- “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)

Wednesday, February 2: AI and Autonomous Systems | Promise & Perils
Note: Your group policy memo is due today!
- “The Experience Machine” by Robert Nozick (Excerpt from Anarchy, State, and Utopia, 1974)
- “AI Now 2019 Report,” pp. 6-24, 36-49 (skim the rest), (AI Now Institute, 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)

Friday, February 4: 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)

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)

Monday, February 7: AI and Autonomous Systems | Bringing to Life
We will be joined in class by: Joshua Browder and Erik Brynjolfsson

Wednesday, February 9: AI and Autonomous Systems | Rights & Responsibilities
- “Artificial Intelligence, Automation and Work” by Daron Acemoglu and Pascual Restrepo, pp. 1-15 (NBER, 2018)
- “AI and the Economy” by Jason Furman and Robert Seamans (NBER, 2018)
- “New World Order” by Erik Brynjolfsson, Andrew McAfee, and Michael Spence (Foreign Affairs, 2014)

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)

Thursday, February 10 (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 11: AI and Autonomous Systems | Making Choices
- “A Basic Income for All” by Philippe Van Parijs (Boston Review, 2000)
- “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)
Monday, February 14: Power of Private Platforms | Promise & Perils

- 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)
- Timothy Garton Ash, Free Speech, “Post-Gutenberg” + ch. 2
- "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
- "Real-Time Politics: The Internet and the Political Process" by Philip Agre (The Information Society, 2002)

Wednesday, February 16: Power of Private Platforms | Technical Deep Dive

Note: Your philosophy paper is due today!

- "Exposure to ideologically diverse news and opinion on Facebook" by Eytan Bakshy, Solomon Messing, Lada A. Adamic (Science, 2015)

Supplementary:
- "Personalized News Recommendation Based on Click Behavior" by Jiahui Liu, Peter Dolan, and Elin Ronby Pedersen (Google, 2009)
- "Filter Bubbles, Echo Chamber, and Online News Consumption" by Seth Flaxman, Sharad Goel, and Justin M. Rao (Public Opinion Quarterly, 2016)

Thursday, February 17 (SECTION): Power of Private Platforms | Tensions & Trade-offs

- Case Study: Platforms
- "The Social Responsibility of Business is to Increase its Profits" by Milton Friedman (The New York Times Magazine, 1970)
- Individual profile: Minnie Ingersoll

Friday, February 18: Power of Private Platforms | Bringing to Life
We will be joined in class by Kara Swisher and Prof. Alex Stamos

Wednesday, February 23: Power of Private Platforms | Rights & Responsibilities

- "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)
- "Tech Platforms and the Knowledge Problem" by Frank Pasquale (American Affairs, 2018)

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)
Friday, February 25: Power of Private Platforms | Making Choices

Note: Technical Assignment #2 is due on Sunday!

- “Free Speech Is Not the Same As Free Reach” by Renee DiResta (WIRED, 2018)

Supplementary:

Monday, February 28: Technology, Inclusion, and Inequality | Promise & Perils

- Iris Marion Young’s The Five Faces of Oppression (2014)
- “When diversity delivers—and when it doesn’t” Interview with Scott Page (2017)
- “Discriminating Systems: Gender, Race, and Power in AI” by Sarah West, Meredith Whittaker, and Kate Crawford (AI Now Institute, 2019)
- “What Really Happened when Google Ousted Timnit Gebru” by Tom Simonite

Supplementary:
- Charles Taylor’s The Politics of Recognition (New Contexts of Canadian Criticism, 1997). A seminal work on the topic; a rich philosophical argument
- Groups of diverse problem solvers can outperform groups of high-ability problem solvers (Proceedings of the National Academy of Sciences, 2004)
- “Advancing Racial Literacy in Tech” by Jesse Daniels, Mutale Nkonde, Darakhshan Mir (Data and Society, 2019)

Wednesday, March 2: Technology, Inclusion, and Inequality | Technical Deep Dive

Note [WIM students only]: Your revised philosophy paper is due today!

- Unlocking the Clubhouse: The Carnegie Mellon Experience by Margolis and Fisher
- Tech Leavers Study (supported by Kapor Center for Social Impact and Ford Foundation) https://www.kaporcenter.org/tech-leavers/
- “Reflecting On One Very, Very Strange Year At Uber” by Susan Fowler (Susan’s blog, 2017)

Supplementary:
- Women and Minorities in Tech, By the Numbers by Myers (https://www.wired.com/story/computer-science-graduates-diversity/)
- Another site that has statistics on tech company diversity (in comparison to the US population) is: Diversity in Tech — Information is Beautiful
• **A Large-Scale Quantitative Study of Women in Computer Science at Stanford University** by Redmond, Evans, and Sahami

**Thursday, March 3 (SECTION): Technology, Inclusion, and Inequality | Tensions & Trade-offs**

- Case Study: Technology, Inclusion, and Inequality

**Friday, March 4: Technology, Inclusion, and Inequality | Rights & Responsibilities**

- **Governing a city of unicorns: technology capital and the urban politics of San Francisco** by Donald McNeill (*Urban Geography*, 2016)
- SF Board of Supervisors report **“Initial Public Offering and Income Inequality”** (2019)
- **East Of Palo Alto’s Eden: Race And The Formation Of Silicon Valley** by Cutler (pp. 24-33, 38-42, 48-68)

Supplementary:

- **Unintended consequences on gender diversity of high-tech growth and labor market polarization** by Elsie Echeverri-Carroll et al. (*Research Policy*, 2018) (pp. 209-212, 216)
- **“Rich corporations, poor societies: The financialisation of Apple”** by Rodrigo Fernandez and Reijer Hendrikse (Oct. 2015) (pp. 10-20)
- **The Fall of the Labor Share and the Rise of Superstar Firms** by David Autor, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reen (*Quarterly Journal of Economics*, 2019)
- **The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power** by Pawel Popiel (*Communication, Culture and Critique*, 2018) (pp. 578-580)
- **How Silicon Valley Broke the Economy** by Adrian Chen (*The Nation*, 2019)

**Monday, March 7: Technology, Inclusion, and Inequality | Making Choices**

- **“Curbing the New Corporate Power”** by K. Sabeel Rahman (Boston Review, 2015)
- **“Sources of Tech Platform Power”** by Lina Khan (Georgetown Tech Law Review, 2018)
- Human Rights Watch: **Big Tech's Heavy Hand Around the Globe**
- **“Combatting Anti-Blackness in the AI Community”** by Devin Guillory (arXiv, 2020)

Supplementary:

- **2018 World Inequality Report** coordinated by Facundo Alvaredo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (pp. 5-16)
- The New Republic: **The Silicon Valley Economy Is Here. And It's a Nightmare** by Lia Russell
- **“Here’s How We Can Break Up Big Tech”** by Elizabeth Warren (2019)

**Wednesday, March 9: Technology, Inclusion, and Inequality | Bringing to Life**

*We will be joined in class by Tracy Chou and Lily Gangas*
● “We’re the Organizers of the Google Walkout. Here Are Our Demands” by Claire Stapleton, Tanuja Gupta, Meredith Whittaker, Celie O'Neil-Hart, Stephanie Parker, Erica Anderson, and Amr Gaber (The Cut, 2018)
● “When Does Activism Become Powerful?” by Hahrie Han (New York Times, 2019)
● Individual profile: Sam King

Friday, March 11: Wrapping Up
Note: Your Final Reflection Paper is due today!
● “The Ones Who Walk Away from Omelas” by Ursula K. Le Guin (1973)
● “With Great Tech Comes Great Responsibility” by Mozilla Foundation (2020)

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)