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 1:00-2:20PM PT
Course Website: https://web.stanford.edu/class/cs182/

<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 Assistants 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 Assistants.

Hilary Cohen (hilco), Head TA
Elena Goldstein (egold21)
Anshul Gupta (agupta21)
Will Kenney (wkenney)
Allison Lettiere (agl12)
Hannah Mieczkowski (hnmiecz)
Luke Miller (lukedm)
Prach Panchakunathom (prachpan)
Natasha Patel (natashap)
Jeff Sheng (jtsheng)
Chloe Stowell (stowell)
Luke Sturm (lsturm)
Anastasiya Vitko (avitko)

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 the Course Manager, Ananya Karthik (ananya23@stanford.edu) or Head TA, Hilary Cohen (hilco@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 is 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 CS181 continues to be taught regularly in a smaller, discussion-focused format.

Targeting students from across multiple disciplines, the course offers 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><strong>Algorithmic Decision-Making</strong></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><strong>Data Collection, Privacy, and Civil Liberties</strong></td>
<td>● Data aggregation, matching, and de-anonymization strategies</td>
</tr>
<tr>
<td></td>
<td>● Facial recognition technology (by public and private actors)</td>
</tr>
<tr>
<td></td>
<td>● Changing norms and laws around privacy across time and cultures (US, China, Europe), including how people balance privacy vs. other goals</td>
</tr>
<tr>
<td></td>
<td>● Consent for different types and uses of data, as well as the idea of companies as “information fiduciaries”</td>
</tr>
</tbody>
</table>
| 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 assistants (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>
</table>
| **Promise and Perils**  
*intro to topic and competing values at stake* | **Bringing to Life**  
*moderated discussion with experts* |
| **Technical Deep Dive**  
*overview of relevant computer science concepts* | **Tensions and Trade-offs**  
*interactive discussion on a case study related to the unit (taking place in SECTION)* |
| **Rights and Responsibilities**  
*policy implications and social science research* | **Making Choices**  
*designing a product/system/policy in light of competing values and trade-offs* |

**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 29
- Group policy assignment on privacy – **due February 10**
- Philosophy paper on autonomous systems – **due February 25**
  - [NOTE: WIM students will have an additional revision **due March 10**]
- Technical assignment on platforms and social networks – **due March 7**
- Final reflection paper — **due on March 19**

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/jpqbarn](http://tinyurl.com/jpqbarn) 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). Section sign-ups will take place online in the first few days of classes. The sign-up form will be available until 5pm on Sunday, January 17th. Please note that section sign-ups are not first-come-first-serve; all submissions in by the sign-up deadline will be treated similarly. After a matching process, your section assignments will be emailed to you. Sections will start the second week of classes. Note that you should only sign up for sections at the CS182 website (you should **not** sign-up for sections on Axess; you should only be signed up for the class lecture on Axess).

**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 be interactive, adequate preparation, willingness to contribute, and capacity for empathetic listening
are all required. You are also required to attend a section every week, and to have read the assigned case in advance. 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.

**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 Teaching 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.

For each lecture, a subset of students (40-50) will be "empaneled," meaning they will be joining a smaller Zoom room that allows for more direct interaction with the professors. **You will be on a panel approximately four times during the quarter.** We intend to make our lectures highly interactive, and empaneled students will be expected to engage fully in class discussion. Students who are not empaneled on a given day will also be able to see and hear interactions between the professors and the empaneled students, as well as to ask questions via Zoom chat (or potentially another question-asking platform). We will let students know in advance which days they will be empaneled, and they will have a dedicated Zoom link to join on those days. **Note: Attendance will be taken for empaneled students.**

We highly encourage you to have your video cameras on during this class whether or not you are empaneled. We intend to use a variety of features, including Zoom breakout rooms, to ensure a dynamic learning environment and believe having cameras on will allow us to connect more deeply while we are geographically distributed. **Students will be expected to participate in these active sessions whether or not they are empaneled for a given lecture.**

Zoom links and details on which Zoom room to join will be available on the course website.

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.)

**The Honor Code**
Violating the Honor Code is a serious offense, even when the violation is unintentional. The Honor Code is available at: 

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 a member of course staff during office hours.

**FERPA**
Student Record Privacy Policy
[http://studentaffairs.stanford.edu/registrar/students/ferpa](http://studentaffairs.stanford.edu/registrar/students/ferpa)
**Additional Resources for Learning**

Students who may need academic accommodation based on the impact of a disability should 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.

**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.