CS181: Computers, Ethics, and Public Policy
Winter 2019 M/W/F 1:30-2:50 PM
Cubberley Auditorium

Course Website: cs181.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>
<tr>
<td>Office Hours: Tuesdays from 3-5pm, 420. Sign-up by emailing Pam Goodman at <a href="mailto:pgoodman@stanford.edu">pgoodman@stanford.edu</a> (who will provide location)</td>
<td>Office Hours: Wednesdays from 3-5pm in Gates 180 (drop-in, no sign-up required in advance)</td>
<td>Office Hours: Mondays from 4-6pm in Encina Hall West, 415. Sign-up is online at: <a href="https://www.wejoinin.com/jweinst@stanford.edu">https://www.wejoinin.com/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
Noah Arthurs (narthurs)
Harper Carroll (hcarroll)
Richard Diehl Martinez (rdm)
Elena Goldstein (egold21)
Esther Goldstein (egolds)

Julia Greenberg (jgreenb)
Zach Harned (zharmed)
Janna Huang (jhuang7)
Sunny Kang (seonkang)
Gabriel Karger (gkarger)
Ale Lynberg (alynberg)

Emma Pierson (emmap1)
Daniel Slate (d slate)
Ashwin Sreenivas (ashwinsr)
Chloe Stowell (stowell)
Glenn Yu (glennyu)
Catherine Yuh (cyuh)

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, start with Hilary Cohen, the Head TA (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 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 and Steve Cooper (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 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>
<tbody>
<tr>
<td><strong>Algorithmic Decision-Making</strong>&lt;br&gt;Jan 11-Jan 25</td>
<td>• Use of predictive algorithms in public vs. private settings, with emphasis on the criminal justice system  &lt;br&gt;• Comparative approaches to algorithmic transparency and accountability (e.g., auditing, technological due process)  &lt;br&gt;• Trade-offs between predictive accuracy and competing values (e.g., fairness, transparency, explainability)  &lt;br&gt;• Mapping automated systems onto a society characterized by human judgments, informal norms, and formal rules</td>
</tr>
<tr>
<td>Topic</td>
<td>Overview</td>
</tr>
<tr>
<td>-----------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Data Collection, Privacy, and Civil Liberties</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>
<tr>
<td>Artificial Intelligence and Autonomous Systems</td>
<td>• Aggregate and distributional consequences of automation (e.g., on labor markets, inequality)</td>
</tr>
<tr>
<td></td>
<td>• Role of policy in shaping the impact of AI and automation broadly, as well as specific policy responses (e.g., universal basic income)</td>
</tr>
<tr>
<td></td>
<td>• Autonomous weapons, including discussion of AI for military use, and autonomous vehicles</td>
</tr>
<tr>
<td></td>
<td>• Competitive global landscape of AI development, comparative approaches to governance, and historical precedents for responses (e.g., Asilomar conference)</td>
</tr>
<tr>
<td>Power of Private Platforms</td>
<td>• 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</td>
</tr>
<tr>
<td></td>
<td>• 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)</td>
</tr>
<tr>
<td></td>
<td>• 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, and the creation of “information bubbles”</td>
</tr>
<tr>
<td></td>
<td>• Comparative analysis of how global platforms operate in diverse communities with different speech traditions and politics</td>
</tr>
</tbody>
</table>

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 four units, we will have a two-week 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 technical exercises, policy memos, and philosophical analyses.

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Week 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promise and Perils</td>
<td>Bringing to Life</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>Technical Deep Dive</td>
<td>Tensions and Trade-offs</td>
</tr>
<tr>
<td><em>overview of relevant computer science concepts</em></td>
<td><em>interactive discussion on difficult debates/issues</em></td>
</tr>
<tr>
<td>Rights and Responsibilities</td>
<td>Making Choices</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 and Coursepack**

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. In order to properly respect copyright rules, **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 six 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 four assignments and a final exam. 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 25
- Philosophy paper on privacy – due February 8 [NOTE: WIM students will have an additional revision due later]
- Group policy assignment on autonomous vehicles – due February 27
- Technical assignment on platforms and social networks – due March 15
- Final exam — held on March 20

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 – 10%</td>
<td>Technical Assignment 1 – 10%</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 – 10%</td>
<td>Technical Assignment 2 – 10%</td>
</tr>
<tr>
<td>Final Exam – 20%</td>
<td>Final Exam – 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; Old Union Room 207). 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. Section sign-ups will take place online between 5:00pm Thursday, January 10th and 5:00pm Sunday, January 13th. The sign-up form will be available on the CS181 web site (cs181.stanford.edu). After a matching process, your section assignments will be e-mailed out to you by Wednesday, January 16th. Sections begin the second week of classes (i.e., next week). Note that you should only sign up for sections at the CS181 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 Sunday at 5pm are treated similarly.
Class Participation

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 come to 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.

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 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.)

**Extra credit.** There will be given an optional extra credit assignment, adding up to 5 points to your final grade, which will be due by March 10. Details will be offered in week 4 of the class.

**Attendance at lectures and sections is mandatory.** If a student has a prolonged illness, varsity athletic competitions, 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 (such as varsity athletic competitions or prolonged illness), 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.

**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](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
563 Salvatierra Walk
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/