Ethics of Technology: 24.131

instructor: Milo Phillips-Brown / milopb@mit.edu
updated from version I taught in Fall 2019

Course description
This class introduces philosophical ethics through contemporary issues concerning technology. Topics include: algorithmic bias and fairness, the singularity, privacy, surveillance, and free speech. A running theme will be how technologies are situated within social and political systems, and what kinds of ethical implications that brings.

Grading and assessment
You learn philosophy by doing philosophy: it's vital that you come to class, participate, and do the readings, which are mandatory.

- Attendance and participation   35%
- In-class quizzes   10%
- Paper 1: 1,000 words   10%
- Group assignment: Algorithmic bias report   10%
- Paper 2: 1,500 words   15%
- Paper 3 (topic of your choosing): 2,000 words   20%

Missed classes and late assignments are excused only with a note from Student Support Services. You can make up for a missed class or a day-late assignment by attending a talk related to the ethics of technology and writing a 250-word reflection.

About participation
You should expect to disagree with the authors we read and with others in class, me included. Whatever the disagreement, always be respectful. Think of the class as a team: we're trying to learn together. Also, talking in class may not be for you (although I'd encourage you to try!). If that's so, please email me with thoughts or questions—even a few sentences is great—and I will email back. Emailing will boost your participation grade and help you learn. You can also increase your participation grade by attending talks related to the ethics of technology and writing a 250-word reflection.

Academic integrity
Plagiarism and other forms of academic dishonesty will not be tolerated. If you're in doubt about what counts as plagiarism, consult me or the Academic Integrity Handbook. Other forms of academic dishonesty include: double submission of papers, falsification of records, and aiding dishonesty. If academic dishonesty is proven, there will be severe disciplinary consequences, which may include any or all of: a redo of the assignment for a reduced grade; failure on the assignment; failure of the course; a letter in the student's file in the Office of Student Citizenship or referral to the Committee on Discipline. If you're tempted to plagiarize, it's always better to speak to me, your advisor,

academic deans, or another trusted authority on campus who can help you.

Accommodation

If you have a disability, you have a right under the American with Disabilities Act to an accommodation.To request an accommodation, contact [Student Disabilities Services](). If you are eligible, they will give you an accommodation letter that you should bring to me as soon as possible. That way you and I can work together to make sure all of the course content is accessible to you. I take this very seriously.

Electronics

Electronics *are not permitted in class*, unless you have an accommodation or have explained your use of them to me in advance. Using electronics in class has been shown to make it harder for both [users]() and [their classmates]() to learn.

If you have concerns

If you have concerns with anything about the class for any reason, I invite you to tell me anonymously via [this form](). (You can also contact the [Ombuds Office]().) Everyone in the class deserves to feel that they're in a good environment. No issue is too small.

Readings

First day

1. [The problem with AI ethics]() *The Verge* James Vincent (2019)

Morality in the machine

2. [The Moral Machine experiment]() Edmond Awad, et al. (2018)
   [Why the Moral Machine is a monster]() Abby Jaques (manuscript)

3. [Moral relativism]() Philippa Foot (1979)
   [Should a self-driving car kill the baby or the grandma? Depends on where you're from]() *MIT Technology Review* Karen Hao (2018)

4. [Do artifacts have politics?]() Langdon Winner (1980)
   [From park bench to lab bench: what kind of future are we designing?]() (video) *TEDx talk* Ruha Benjamin (2015)

Bias in, bias out

5. [Big data's disparate impact]() part 1, Solon Barocas and Andrew Selbst (2016)
   [The seductive diversion of 'solving' bias in artificial intelligence]() *Medium* Julia Powles and Helen Nissenbaum (2018)

6. [Demarginalizing the intersection of race and sex](#) Kimberlé Crenshaw (1989)
   [Gender shades](#) (video), Joy Buolamwini (2018)

7. ["Night to his day": the social construction of gender](#) Judith Lorber (1994)
   [Missed connections: what search engines say about women](#) *Bitch Magazine* Safiya Noble (2012)

Algorithmic fairness: formal definitions and their limits

8. [Justice as fairness](#) John Rawls (1958)
   [Machine bias](#) *ProPublica* Julia Angwin, et al. (2016)

9. [Fairness in machine learning: against false positive rate equality as a measure of fairness](#) Robert Long (manuscript)
   [Locking up our own](#) introduction, James Forman Jr. (2017)

10. [Measuring algorithmic fairness](#) part 2, Deborah Hellman (2019)
    [To predict and serve](#) Kristian Lum and William Isaac (2016)

Surveillance and privacy

11. [Panopticonism (in Discipline and punish)](#), Michel Foucault (1975)
    [Stuck In China's panopticon](#) (podcast) *NPR*'s *Planet Money* Alisa Chang (2019)

12. [Big other: surveillance capitalism and the prospects of an information civilization](#) Shoshana Zuboff (2015)
    [I'm being watched at Amazon Go—and I don't care](#) *Axios* Erica Pandey (2019)

13. [Contextual integrity up and down the data food chain](#) Helen Nissenbaum (2018)
    [What's NSA surveillance got to do with me?](#) (video) *ACLU* (2013)

What does technology have to do with disability?

14. [Causing disability, valuing disability](#) Elizabeth Barnes (2014)
    [Deep sea diving... in a wheelchair](#) (video) *TED talk*, Sue Austin (2013)

15. [Capabilities and disability: the capabilities framework and the social model of disability](#) Tania Burchardt (2010)
    [Deaf culture and cochlear implants: genocide or salvation?](#) *Time* Lydia Denworth (2014)

16. [Sociotechnical considerations for accessible visualization design](#) Crystal Lee, et al. (2019)
    [Designing UX for people with disabilities](#) *UX Studio* Alexandra Kovács (2018)

Free speech and democracy

17. [On liberty](#) chapter 1, John Stuart Mill (1859)
    [Facebook and the first amendment: policing free speech on the platform](#) (podcast) *WBUR's On Point* Meghna Chakrabarti et. al (2019)

18. [The epistemology of democracy](#) Elizabeth Anderson (2006)
    [Letter to Facebook employees on political ads](#) Mark Zuckerberg (2019)

19. [It's the (democracy-poisoning) golden age of free speech](#) *Wired* Zeynep Tufekci (2018)
    [Facebook finally has a good idea: the company's latest plan to police toxic social media content is intriguing—and even laudable.](#) *New York Times* Kara Swisher (2019)

The robot apocalypse

20. [The singularity: a philosophical analysis](#) sections 1, 2 and 4, David Chalmers (2010)
    [The troubling trajectory of technological singularity](#) *Forbes* Jayshree Pandya (2019)

21. [Existential risk prevention as global priority](#) *skip* section 2, Nick Bostrom (2013)
    [We shouldn't be scared by 'superintelligent A.I.'](#) *New York Times* Melanie Mitchell (2019)

22. [Stop the robot apocalypse](#) *London Review of Books,* Amia Srinivasan (2015)
    [Hello quantum world! Google publishes landmark quantum supremacy claim](#) *Nature News* Elizabeth Gibney (2019)

Employee activism

23. [Civil disobedience](#) Henry David Thoreau (1849)
    [Three years of misery inside Google, the happiest company in tech](#) *Wired* Nitasha Tiku (2019)

24. [Letter from Birmingham jail](#) Martin Luther King (1963)
    [We are Google employees. Google must drop Dragonfly](#) *Medium* Google Employees Against Dragonfly (2018)

25. [The Crito](#) Plato
    [Google is cracking down on its employees' political speech at work](#) *Vox* Shirin Ghaffary (2019)