CS 492: Social Implications of Computing: Winter 2023 syllabus

Sections 003 and 004: Discrimination, Privacy, and Surveillance in Computing

1. Contact information:
    a. Instructors:
    Daniel G. Brown, Ph.D.
    David R. Cheriton School of Computer Science
    Davis Centre 2516
    200 University Avenue West
    Waterloo, ON N2L 3G1
    dan.brown@uwaterloo.ca
    1-519-888-4567, ext. 36278
    Office hours: Mondays, 10-11 am, DC 2516

    Maura R. Grossman, J.D., Ph.D.
    David R. Cheriton School of Computer Science
    Davis Centre 2501
    200 University Avenue West
    Waterloo, ON N2L 3G1
    maura.grossman@uwaterloo.ca
    1-519-888-4567, ext. 37522 (office)
    1-212-861-8097 (voicemail)
    Office hours: by request; email to schedule a time
    b. Teaching Assistants:
    Ben Armstrong
    Liam Hebert
    Chris Liscio
2. Course description:
    a. CS 492 as a whole:
    This course considers the impact of computing on individuals, organizations, and society. Students learn how to make informed judgements about the social and ethical consequences of deploying computing technologies and how to foster critical thought and professional responsibility.
    b. This offering of CS 492:
    This course will consider the ways in which computing systems may enable, encourage, or prevent societal discrimination of a variety of types. Students will examine how algorithms and artificial intelligence systems may provide different outcomes for different populations, and what legal, technical, and societal interventions can ameliorate harms caused by such algorithms and systems.
3. Course goals:
    a. CS 492 as a whole:

At the end of the course, students should be able to:

- Identify important social implications of computerization
- Describe how to behave ethically as a computer scientist
- Propose technological and social/policy solutions to current challenges presented by computerization (recommended actions)
- Combine their opinions with information from existing research in both written and oral formats
- Collaborate with peers on open-ended topics where there is often no clear definitive response
- Formulate clear, constructive feedback on arguments constructed by their peers
- Identify where their knowledge of computer science provides a valuable contribution to the overall dialogue about the social and ethical implications of computing

b. This offering of CS 492:
Our primary goal is for students in technology careers to be able to (i) discuss with colleagues and others the positive or negative discriminatory consequences of technological innovations, and (ii) explain techniques to remove bias from data collections, computer algorithms, or artificial intelligence applications, and why they might or might not work in specific circumstances.

To be specific, our objectives are that at the end of this course, students will be able to:

- Describe, using historical context and a sociocultural lens, how computing systems can either deliberately or systematically aid and abet discrimination among human populations;
- Reason effectively with colleagues and others about the positive or negative discriminatory consequences of technological innovations;
- Explain techniques available to remove bias from data collections, computer algorithms, or artificial intelligence applications, and why they might or might not work in specific circumstances; and
- Describe how legal or policy changes at a broad level may or may not improve the situation of computing and discrimination in specific cases.

4. Course meetings and logistics:
We will be teaching two sections of CS 492 in Winter 2023:

- Section 003 will meet face-to-face from 10:30 am-12:20 pm on Tuesdays, in MC 4060.
- Section 004 will meet fully online, synchronously, from 2:30-4:20 on Tuesdays, via this Zoom link: https://uwaterloo.zoom.us/j/98279309244?pwd=QUlwT0hLM0wrL1JmK2F4akgrdVlWQT09
- Students must remain in their assigned section.

Class discussion will comprise a significant component of the course, so attendance at all class sessions is required unless absences are excused by the instructors in advance. An excessive number of absences will impact the participation component of a student's grade. Please do not attend class if you are ill and in self-isolation; just let the instructors know you are unable to attend class. Class sessions will not be recorded. There is a Slack channel for discussions in this course, where topics for essays, announcements, and other information will be shared: XXXXX

5. From the news:
   Most weeks, CS 492 will start with a "from the news" segment featuring a relevant article from recent news. We will supply you with the reading or video (which will be short) on the Friday before the course meeting. You will find it on the "from-the-news" channel on the course Slack.
   Please submit articles or videos for us to consider for the "from-the-news" segment: post them in the "from-the-news" channel on the course Slack by Wednesday evening the week before the class you want to discuss the article or video , so that we have time to review it. Tell us why you think it would be a good choice. Submitting articles or videos in this way will count towards the it "participation" component of the course grade. If we select your from-the-news submission in a week, we will give you a 1% bonus mark towards your final grade in the course.
6. Course readings:
   There will be no textbook for this course. All course materials are attached as links to this course syllabus, are freely available online, or (if subject to copyright protection) are available to students registered for this course via the LEARN link for the course, https://learn.uwaterloo.ca/d2l/home/869391
7. Course evaluation and grading:
   Evaluation in this course will include three components:
   a. Writing assignments: 45%
      - There will be weekly writing assignments, which will be commentaries on the readings. There will be around eleven such commentaries assigned; only the top six scores will count towards a student's grade. Students will receive regular, written feedback on their papers from one of the teaching assistants. Commentaries will have a prompt given every week at the end of class, and will also be posted online in the Slack channel. They will be due every week on Thursday at 11:59 PM via the Dropbox connected to the course via LEARN, in either .doc/.docx format or as a PDF.
      - The six (6) best commentaries or position papers submitted by a student will each count for 7.5% of the student's final grade. Therefore, the written component of the course will account for a total of 45% of a student's final grade. Should a student wish, they may submit any commentary to an instructor by email, in an editable word-processing format (i.e., .doc or .docx) for additional feedback (beyond that received from the teaching assistant), particularly with respect to improving their writing style.

- The University of Waterloo's Writing and Communication Centre (WCC) works across all faculties to help students clarify their ideas, develop their voices, and write in the styles appropriate to their disciplines. Centre staff offer one-on-one support in planning assignments and presentation, using and documenting research, organizing and structuring papers, and revising for clarity and coherence. We particularly encourage using the WCC for students whose English or writing skills are not at the level where they wish them to be.
- Your weekly writing assignments should be approximately two pages (1000 words) long.
- These assignments will be marked on a scale from zero to five. The weekly topic will be announced at the end of each class and will also be posted to the course Slack channel.
- We have some examples of good essays for previous offerings of this course and related courses available on the LEARN page for this course.

b. Class participation: 20%
    - We will assess students' participation on thoughtfulness and quality, rather than quantity. What we are looking for is for students who can demonstrate meaningful engagement with the material, justify their arguments, and respond respectfully to contrary positions presented by others. During the course of the semester, students are likely to be asked to advance or defend (either orally or in writing or both) positions that may differ from their own personal views.
    - Suggesting articles or videos for the "from-the-news" component of the class will count towards a student's class participation grade, as will contributing meaningfully to discussions in the course Slack channel.
    - Students for whom English communication is challenging may need to work especially hard for this component of the course. Participation in the Slack channel or the Zoom chat thread (for the online version of the course) are particularly encouraged. Please let us know how we can help you participate if English is not your first language or you are introverted (for example, should we call on you?)
c. Group project: 35%
    - There will be a small-group project that will involve a small team comprised of two to four students. Each team will submit a joint work product that may consist of a creative project, such as the development of an online program or tool, an application, a chatbot, a website, a video presentation, or a podcast.
    - We urge you not to do a podcast for this project unless you have something novel to say or a particularly unique style to present the topic; it is too easy to just make a chatty conversation with not much content, which will not garner a good grade. If you do a podcast, please focus on a small number of topics, find articles to discuss that are not in

the course readings, keep a tight thematic focus, and find a very creative way to present the material.

- The group projects will involve a technological issue of the sort addressed in the course, regarding a computing topic that connects to discrimination, surveillance, and/or privacy and how they connect to computing. Teams will identify their topic, confirm the appropriateness of the topic with the instructors, divide the labour among their team members, and accommodate different perspectives, viewpoints, and expertise of the team members.
- If students encounter undue challenges in the small-group project work, they should contact the instructors promptly; in particular, please do not wait until the week before the project is due to indicate that group members have been unresponsive to their assignments.
- Students may, but are not required, to prepare a short presentation to give in the final class of the semester. For students who choose to present, this will contribute to their class participation grade.
- The submission for the project consists of:
    1. A one-page description of the project, including the group members' names, what question was being addressed, what form the project took, and a summary of the findings. This may also include a link to code repositories, or other specific items you want to have considered in your mark.
    2. A five-minute video tour of the project and its findings. This should focus on results and demonstrating any system built. For a podcast project, identify the articles you discussed, how you found them, and the technology you used in the preparation of the podcast.
    3. The actual project: a usable website, a finished podcast, a functional chatbot, a playable game, or whatever the final form takes.

  All of this should be submitted via the course's LEARN dropbox, with the due date of 10 April at 11:59 PM. There will be no extensions for late projects; we need to submit course grades on time!

This course has neither a midterm nor a final examination.

8. Rules for small-group work:

   The project should only be work done by your group members: in particular, nothing you submit for marking should be from anyone else. You may submit other people's work that has been inspiration or that your project seeks to critique. The one exception to this rule is that you can use existing libraries or code bases, with proper acknowledgment; that said, the main intellectual contribution of your small-group project must be your own.

   For the weekly essays, they must be entirely your own writing; you may discuss your thoughts with your classmates through the lectures and in your own conversations outside class, whether on Slack or otherwise. A good rule of thumb is that if the

instructors pull you aside and asks you to explain the arguments underlying your writing, you should be able to explain what you mean by every sentence; if not, you don't understand your own writing, which suggests that it's not sufficiently your own.

9. Marks appeals:
   Should any student wish to discuss the written feedback or their mark on an individual commentary or position paper, that student should first reach out to the teaching assistant who provided the feedback or mark by email to schedule a time to discuss the comments or mark. Should the student continue to be dissatisfied with the feedback or mark provided by the teaching assistant, they should reach out to the instructors by email to schedule a time to discuss their concerns. This should be rare. Formal appeals may result in assignment of the same grade, a higher grade, or a lower grade than the previous grade awarded by the teaching assistant.
10. Course website:
    The readings for CS 492 Sec. 003 and 004 will be available via this link: https://cs.uwaterloo.ca/~browndg/492W23/index.shtml; this link also includes this syllabus, and other materials.
11. Land acknowledgment:
    The University of Waterloo is located on the traditional territory of the Anishinaabeg, Neutral, and Haudenosaunee people. The main campus is located on the Haldimand Tract, which was granted to the Six Nations after the American Revolution. This tract includes six miles of land on either side of the Grand River. Both instructors are non-Indigenous settlers who came to this country to teach and do research, and we are committed to incorporating Indigenous content into my teaching work. At the University of Waterloo, active work toward reconciliation takes a number of forms and is centralized within the Office of Indigenous Relations.
12. COVID continuity plan:
    If the university alters the delivery structure for courses, we will respond as follows:
    For cancellation of in-person classes, we will switch to Zoom-based instruction.
    Students will be made aware of such arrangements via the Slack site for the course.
    If one instructor takes ill, courses will be delivered by the other instructor; if both of us take ill, some classes may need to be delivered by teaching assistants, or via the course being taken fully online during the instructors' isolation period.
    No specific accommodation will be made for students in self-isolation; they can connect with the instructors during office hours or by appointment and can obtain the weekly essay topic via Slack.
    The course has neither a midterm nor a final.
13. Waterloo policies:
    a. Academic Integrity
       In order to maintain a culture of academic integrity, members of the University of Waterloo community are expected to promote honesty, trust, fairness, respect, and responsibility. [Check the Office of Academic Integrity for more information.]
    b. Grievance
       A student who believes that a decision affecting some aspect of their university life has been unfair or unreasonable may have grounds for initiating a grievance. Read

Policy 70, Student Petitions and Grievances, Section 4. When in doubt, please be certain to contact the department's administrative assistant who will provide further assistance.

c. Discipline
A student is expected to know what constitutes academic integrity to avoid committing an academic offence, and to take responsibility for their actions. [Check the Office of Academic Integrity for more information.] A student who is unsure whether an action constitutes an offence, or who needs help in learning how to avoid offences (*e.g.*, plagiarism, cheating) or about "rules" for group work/collaboration should seek guidance from the course teaching assistants, course instructors, academic advisor, or the undergraduate associate dean. For information on categories of offences and types of penalties, students should refer to Policy 71, Student Discipline. For typical penalties, check Guidelines for the Assessment of Penalties.

d. Appeals
A decision made or penalty imposed under Policy 70, Student Petitions and Grievances (other than a petition) or Policy 71, Student Discipline may be appealed if there is a ground. A student who believes they have a ground for an appeal should refer to Policy 72, Student Appeals.

e. Note for Students with Disabilities
AccessAbility Services, located in Needles Hall, Room 1401, collaborates with all academic departments to arrange appropriate accommodations for students with disabilities without compromising the academic integrity of the curriculum. If you require academic accommodations to lessen the impact of your disability, please register with AccessAbility Services at the beginning of each academic term. Please do not wait until the end of the term to advise us you need an accommodation; it is very challenging for us to respond to needs for accommodation late in the term.

f. Mental Health
If you or anyone you know experiences severe academic stress, difficult life events, or feelings of severe anxiety or depression, we strongly encourage you to seek support.
On campus resources include:
- Campus Wellness at https://uwaterloo.ca/campus-wellness/
- Counseling Services at counselling services@uwaterloo.ca / 1-519-888-4567, ext. 32655 / Needles Hall 2nd Floor (NH 2401)
- MATES one-to-one peer support program offered by WUSA and Counselling Services at mates@uwaterloo.ca
- Health Services located across the creek from the Student Life Center (SLC) / 519-888-4096

Off campus resources include:
- Good2Talk (24/7) free confidential helpline for post-secondary students at 1-866-925-5454
- Here 24/7 Mental Health and Crisis Service Team at 1-844-437-3247

- OK2BME support services for gay, lesbian, bisexual, transgender, or those questioning their sexual or gender identity at 1-519-884-0000, ext. 213
- EMPOWER ME 1-833-628-5589 for Canada / U.S.A. / Other countries
  http://studentcare.ca/rte/en/IHaveAPlan_WUSA_EmpowerMe_EmpowerMe
  EMPOWER ME in China: China North 108007142831;
  China South 108001402851

g. Diversity

It is the Faculty of Mathematics' intention that students from diverse backgrounds and perspectives be well served in its courses, and that students' learning needs be addressed both in and out of the classroom. We recognize the immense value of diversity in identities, perspectives, and contributions that students bring, and the benefit it has on the educational environment. Your suggestions are encouraged and appreciated. Please let your instructors and teaching assistants know if there are ways to improve the effectiveness of this course for you personally, or for other students or student groups. In particular:

- We will gladly honour your request to address you by an alternate/preferred name or gender pronoun. Please advise us of this preference early in the semester so that we may make appropriate changes to our records;
- We will honour your religious holidays and celebrations. Please inform us of those at the start of the semester; and
- We will follow AccessAbility Services guidelines and protocols on how best to support students with different learning needs. Please do not wait until the end of the term to advise us that you need an accommodation; it is very challenging for us to respond to needs for accommodation late in the term.

h. Intellectual Property**:** Students should be aware that this course contains the intellectual property of their instructors, TAs, and/or the University of Waterloo. Intellectual property includes items such as:

- Lecture content, spoken and written (and any audio/video recording thereof);
- Lecture handouts, presentations, and other materials prepared for the course (e.g., PowerPoint slides);
- Questions or solution sets from various types of assessments (e.g., assignments, quizzes, tests, final exams); and
- Work protected by copyright (e.g., any work authored by the instructors or TAs or used by the instructors or TAs with permission of the copyright owner).

Course materials and the intellectual property contained therein, are used to enhance a student's educational experience. However, sharing this intellectual property without the intellectual property owner's permission is a violation of intellectual property rights. For this reason, it is necessary to ask the instructors, TAs and/or the University of Waterloo for permission before uploading and sharing the intellectual property of others online (e.g., to an online repository).

Permission from an instructors, TAs or the University is also necessary before sharing the intellectual property of others from completed courses with students taking the same/similar courses in subsequent terms/years. In many cases, instructors might be happy to allow distribution of certain materials. However, doing so without expressed permission is considered a violation of intellectual property rights.

Please alert the instructors if you become aware of intellectual property belonging to others (past or present) circulating, either through the student body or online. The intellectual property rights owner deserves to know (and may have already given their consent).

WEEKLY CLASS SCHEDULE WITH ASSIGNMENTS

1. <u>Tuesday, January 10: Welcome, Introduction, Overview, and Some Case Studies</u>

- *What is this course about?*
  - K. Hao, *Artificial intelligence is creating a new colonial world order,* MIT Technology Review (2022), available at https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/.
  - Read part one of this series on AI colonialism: K. Hao, *South Africa's private surveillance machine is fueling a digital apartheid*, MIT Technology Review (2022), available at https://www.technologyreview.com/2022/04/19/1049996/south-africa-ai-surveillance-digital-apartheid/.
- *A couple of case studies involving discrimination, privacy, and/or surveillance in computing*
  - P. Steinfels, *The Deep Strangeness of the Catholic Church's Latest Scandal*, The Atlantic (2021), available at https://www.theatlantic.com/ideas/archive/2021/08/catholic-priest-jeffrey-burrill-grindr-pillar/619758/.
    - Original article: *Pillar Investigates: USCBB gen sec resigns after misconduct allegations*, The Pillar (2021), available at https://www.pillarcatholic.com/p/pillar-investigates-usccb-gen-sec).
    - How they did it: Z. Davis, *Tabloids, scandal and spying: The U.S. Catholic Church has hit a new, dangerous low point*, America (2021), available at https://www.americamagazine.org/faith/2021/07/22/catholic-church-spying-pillar-burrill-grindr-usccb-privacy-241106).
    - Is this really anything different? Everything old is new again: M. Signorile, *Right-wing Catholic moralists outed a Catholic official as gay*, The Signorile Report (2021), available at https://signorile.substack.com/p/right-wing-catholic-moralists-outed.
    - About the other figures he outed in the '90s: T. Goldstein, *It's Not Up to You*, The New York Times Archive (1993), available at https://www.nytimes.com/1993/06/27/books/it-s-not-up-to-you.html.
    - N. Singer and A. Krolik. *Grindr is fined $11.7 million under European privacy law,* The New York Times (2021), available at https://www.nytimes.com/2021/01/25/business/grindr-gdpr-privacy-fine.html.
  - J. Dastin, *Amazon scraps secret AI recruiting tool that showed bias against women*, Reuters (2018), available at https://www.reuters.com/article/us-amazon-com-jobs-automation-

insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G.

- Whoops, we did it again:
  J. Del Rey, *A leaked Amazon memo may help understand why the tech giant is pushing out so many recruiters*, Vox (2022), available at https://www.vox.com/recode/2022/11/23/23475697/amazon-layoffs-buyouts-recruiters-ai-hiring-software

- Optional, but definitely worth watching if you have Netflix and haven't already seen it is the following documentary on how machine learning algorithms can perpetuate existing racial, gender, and class-based inequities:
  Shalini Kantayya, *Coded Bias*, Netflix (2020), available at https://www.netflix.com/ca/title/81328723.

2. **Tuesday, January 17: More Case Studies and Theoretical Approaches to Technology Ethics and Algorithmic Bias**
   - *More case studies*
     - G. Smith, *High-tech redlining: AI is quietly upgrading institutional racism*, Fast Company (2018), available at https://www.fastcompany.com/90269688/high-tech-redlining-ai-is-quietly-upgrading-institutional-racism
     - C. McIlwain, *Can Technology Help Make Housing Fairer?*, MIT Technology Review (2020). **[NB: This article will be made available through LEARN]**
     - S. Kelley & A. Ovchinnikov, *Anti-discrimination Laws, AI, and Gender Bias: A Case Study in Non-mortgage Fintech Lending*, SSRN (2020), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3719577.
     - M. Kay et al., *Unequal Representation and Gender Stereotypes in Image Search Results for Occupations*, Proceedings of CHI '15 (2015), available at https://dl.acm.org/doi/10.1145/2702123.2702520.
   - *The definition of "discrimination" and an overview of the problem*
     - F.J. Zuidvereen Borgesius, *Discrimination, Artificial Intelligence, and Algorithmic Decision-Making*, Council of Europe (2018), available at https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73.
     - J. Kleinberg et al., *Discrimination in the Age of Algorithms*, Journal of Legal Analysis (2018), available at https://academic.oup.com/jla/article/doi/10.1093/jla/laz001/5476086?login=true.
   - *An introduction to the ethics of AI*
     - J. Chou et al., *In Pursuit of Inclusive AI*, Microsoft Design (2018), available at https://www.microsoft.com/design/assets/inclusive/InclusiveDesign_InclusiveAI.pdf
     - S.A. Friedler, *The (Im)possibility of Fairness: Different Value Systems Require Different Mechanisms For Fair Decision Making*, Communications of the ACM (2021), available at https://dl.acm.org/doi/pdf/10.1145/3433949

3. **Tuesday, January 24: Systemic Racial Discrimination in the Criminal Justice System and Beyond ***Special Guest [Christopher Taylor] will join us*****
   - *The criminal justice system*
     - S. Bastek, an excerpt from C. O'Neil, *Weapons of Math Destruction*, The American Scholar (2016), available at https://theamericanscholar.org/weapons-of-math-destruction/.
       - C. O'Neil, *Civilian Casualties: Justice in the Age of Big Data,* in Weapons of Math Destruction – How Big Data Increases Inequality and Threatens Democracy, ch. 5, pp. 84-104 (2016). **[NB: This chapter will be made available through LEARN]**
     - J. Angwin et al., *Machine Bias*, ProPublica (2016), available at https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
     - K. Hao and J. Stray, *Can you make AI fairer than a judge? Play our courtroom algorithm game,* MIT Technology Review (2019), available at https://www.technologyreview.com/2019/10/17/75285/ai-fairer-than-judge-criminal-risk-assessment-algorithm/.
     - *Watch this video thread of a real case where the use of recidivism assessment tool was questioned by a judge:* https://grossman.uwaterloo.ca/riskassessment.mp4.
     - J.C. Brigham & D.J. Ready, *Own-race bias in lineup construction*, Law and Human Behavior (1985), available at https://link.springer.com/article/10.1007/BF01044480
   - *Facial recognition*
     - J. Buolamwini and T. Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, Proceedings of Machine Learning Research (2018), available at https://proceedings.mlr.press/v81/buolamwini18a.html.
     - K. Hill, *Accused of Cheating By An Algorithm,* The New York Times (2022), available at https://www.nytimes.com/2022/05/27/technology/college-students-cheating-software-honorlock.html.
   - *Healthcare*
     - H. Ledford, *Millions of black people affected by racial bias in health-care algorithms*, Nature (2019), available at https://www.nature.com/articles/d41586-019-03228-6.
     - The original study can be found here: Z. Obermeyer et al., *Dissecting racial bias in an algorithm used to manage the health of populations*, Science (2019), available at https://www.science.org/doi/10.1126/science.aax2342.

4. **Tuesday, January 31: Discrimination Against Indigenous Persons, from Design to Access to Technology**
   - *"Real names" policies and social networks*

- E. Dwoskin et al., *Facebook's race-blind practices around hate speech came at the expense of Black users, new documents show*, The Washington Post (2021), available at https://www.washingtonpost.com/technology/2021/11/21/facebook-algorithm-biased-race/.
- A. Holpunch, *Facebook still suspending Native Americans over 'real name' policy*, The Guardian (2015), available at https://www.theguardian.com/technology/2015/feb/16/facebook-real-name-policy-suspends-native-americans.

- *Search engines and Indigenous languages*
  - *What If You Couldn't Search the Internet in Your Own Language*, Kick Point (2021), available at https://kickpoint.ca/what-if-you-couldnt-search-the-internet-in-your-own-language/.
- *First Nations and Internet access*
  - S.S.F. Young, *Broadband Bruce: Fighting Canada's Digital Divide*, YouTube (2017), available at https://www.youtube.com/watch?v=lJUZcMjTOVI.
  - M. Buell, *Indigenous communities must have internet access on their terms*, Toronto Star (2021), available at https://www.thestar.com/opinion/contributors/2021/01/18/indigenous-communities-must-have-internet-access-on-their-terms.html.
- *AI Colonialism*
  - K. Hao, *A new vision of artificial intelligence for the people,* MIT Technology Review (2022), available at https://www.technologyreview.com/2022/04/22/1050394/artificial-intelligence-for-the-people/.
- *Indigenous-Centred Approaches to Design*
  - M. Fancy, *Designing ethical AI through Indigenous-centred approaches, University of Ottawa* (2020), available at https://techlaw.uottawa.ca/aisociety/news/designing-ethical-ai-indigenous-centred-approaches

5. **Tuesday, February 7: Gender Discrimination and Violence against Women ***Special Guest TK Pritchard will join us*****
   - *The world is literally designed for men*
     - C. Criado Perez, *The deadly truth about a world built for men – from stab vests to car crashes*, The Guardian (2019), available at https://www.atheguardian.com/lifeandstyle/2019/feb/23/truth-world-built-for-men-car-crashes.
   - *… But what if it wasn't*
     - S. Fonseca, *Designing forms for gender diversity and inclusion,* Medium (2017), available at https://uxdesign.cc/designing-forms-for-gender-diversity-and-inclusion-d8194cf1f51.
   - *Abortion access and privacy concerns*

- L. H. Newman, *The Surveillance State Is Primed for Criminalized Abortion,* WIRED (2022), available at https://www.wired.com/story/surveillance-police-roe-v-wade-abortion/.
- J. Famularo and R. Wong, *How the tech sector can protect personal data post-Roe,* Brookings Institution TechTank, available at https://www.brookings.edu/techstream/how-tech-firms-can-protect-personal-data-after-roe-us-privacy-abortion-surveillance/.

- *Bias in NLP systems*
  - T Buonocore, *Man is to Doctor as Woman is to Nurse: The Gender Bias of Word Embeddings*, towards data science (2019), available at https://towardsdatascience.com/gender-bias-word-embeddings-76d9806a0e17.
- *Why are virtually all the AI assistants and bots women?*
  - C. Chin & M. Robinson, *How AI bots and voice assistants reinforce gender bias*, Brookings Institution Report (2020), available at https://www.brookings.edu/research/how-ai-bots-and-voice-assistants-reinforce-gender-bias/.
- *Technology-facilitated gender-based violence ("TFGBV")*
  - S. Dunn, *Technology-Facilitated Gender-Based Violence: An Overview*, Center for International Governance Innovation (2020), available at https://www.cigionline.org/publications/technology-facilitated-gender-based-violence-overview/.
  - J. Clayton and J. Dyer, *Apple AirTags - A perfect tool for stalking,* BBC News (2022), available at https://www.bbc.com/news/technology-60004257.

6. **Tuesday, February 14: Big Data, the Right to Privacy, and the Right to be Forgotten**
   - *Big data*
     - M. Madden et al., Privacy, *Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans*, Washington Law Review (2017), available at https://openscholarship.wustl.edu/law_lawreview/vol95/iss1/6/. → **Read only Introduction, Part I, and Part III at pages 53-67 and 79-95.**
     - J. Vincent, *Google's DeepMind and UK hospitals made illegal deal for health data, says watchdog*, The Verge (2017), available at https://www.theverge.com/2017/7/3/15900670/google-deepmind-royal-free-2015-data-deal-ico-ruling-illegal.
     - S. Shead, *Britain gave Palantir access to sensitive medical records of Covid-19 patients in £1 deal*, CNBC (2020), available at https://www.cnbc.com/2020/06/08/palantir-nhs-covid-19-data.html.
   - *Protecting privacy . . . or no*
     - C.F. Kelly, *Protecting privacy in an AI-driven world*, Brookings Institution Report (2020), available at https://www.brookings.edu/research/protecting-privacy-in-an-ai-driven-world/.
     - C. Dwork and D.K. Mulligan*, It's Not Privacy, and It's Not Fair*, Stanford Law Review (2013), available at

https://www.stanfordlawreview.org/online/privacy-and-big-data-its-not-privacy-and-its-not-fair/.

- *The right to be forgotten . . . or not*
  - S. Kulk and F. Zuidvereen Borgesius, *Privacy, Freedom of Expression, and the Right to Be Forgotten in Europe*, in The Cambridge Handbook of Consumer Privacy (Cambridge Univ. Press 2018), available at https://dare.uva.nl/search?identifier=f7d0f415-3404-426a-8833-12861fee7112.
  - D. Solove, *Is the Right to Be Forgotten Good or Bad? This is the Wrong Question*, Privacy + Security Blog (2014), available at https://teachprivacy.com/right-forgotten-good-bad-wrong-question/.
  - D. Gillmor, *The 'right to be forgotten' doesn't mean we should be censoring Google results*, The Guardian (2014), available at https://www.theguardian.com/commentisfree/2014/jun/24/right-to-be-forgotten-censoring-google-results.
  - E. Dans, *It's Time to Forget the Right to Be Forgotten*, Forbes (2019), available at https://www.forbes.com/sites/enriquedans/2019/09/25/its-time-to-forget-the-right-to-be-forgotten/?sh=4ef8042f715e.
  - B. Auxier, *Most Americans support right to have some personal information removed from online searches*, Pew Research Center (2020), available at https://www.pewresearch.org/fact-tank/2020/01/27/most-americans-support-right-to-have-some-personal-info-removed-from-online-searches/.
- *Case study:* On Our Backs ("OOB")
  - T. Robertson, *Digitization: just because you can, doesn't mean you should*, Tara Robertson Consulting (2016) available at https://tararobertson.ca/2016/oob
  - T. Robertson, *Update on Our Backs and Reveal Digital*, Tara Roberts Consulting (2016), *available at* https://tararobertson.ca/2016/oob-update/

*****NO CLASS ON TUESDAY, FEBRUARY 21: READING WEEK*****

Use this opportunity to catch up on any readings you have not yet completed or to read ahead!

7. **Tuesday, February 28: Discrimination on the Basis of Sexual Orientation and Gender Identity**
   - *Historical roots of facial recognition*
     - B. Aguera y Arcas et al, *Physiognomy's New Clothes*, Medium (2017), available at https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a.
   - *Facial recognition and sexual orientation / gender identity*
     - S. Levin, *New AI can guess whether you're gay or straight from a photograph*, The Guardian (2017), available at https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph.

        - Y. Wang and M. Kosinski, *Deep neural networks are more accurate than humans at detecting sexual orientation from facial images*, Journal of Personality and Social Psychology (2018), available at https://psycnet.apa.org/doiLanding?doi=10.1037%2Fpspa0000098.
        - *But see* B. Aguera y Arcas et al., *Do algorithms reveal sexual orientation or just expose our stereotypes?*, Medium (2018) available at https://medium.com/@blaisea/do-algorithms-reveal-sexual-orientation-or-just-expose-our-stereotypes-d998fafdf477.
        - S. Melendez, *Uber driver raises concerns about transgender face recognition*, Fast Company (2018), available at https://www.fastcompany.com/90216258/uber-face-recognition-tool-has-locked-out-some-transgender-drivers.
        - J. Vincent, *Transgender YouTubers had their videos grabbed to train facial recognition software*, The Verge (2017), available at https://www.theverge.com/2017/8/22/16180080/transgender-youtubers-ai-facial-recognition-dataset.
    - *Deepfakes, AI-generated porn, and women's bodies*
        - S. Cole, Deepfakes *Were Created As a Way to Own Women's Bodies—We Can't Forget That,* Vice (2018), available at https://www.vice.com/en/article/j5kk9d/deepfakes-were-created-as-a-way-to-own-womens-bodieswe-cant-forget-that-v25n2.
        - K. Wiggers, Meet Unstable Diffusion, the group trying to monetize AI porn generators, TechCrunch+ (2022), available at https://techcrunch.com/2022/11/17/meet-unstable-diffusion-the-group-trying-to-monetize-ai-porn-generators

8. **Tuesday, March 7: Discrimination Against Disabled Persons**
    - *Life as a disabled person*
        - P. Smith & L. Smith, *Artificial intelligence and disability: too much promise, yet too little substance*, AI and Ethics (2021), available at https://link.springer.com/article/10.1007/s43681-020-00004-5.
    - *Disability discrimination or accessibility . . . by default*
        - A. Engler, *For some employment algorithms, disability discrimination by default*, Brookings Institution TechTank (2019), available at https://www.brookings.edu/blog/techtank/2019/10/31/for-some-employment-algorithms-disability-discrimination-by-default/.
        - B. Hutchinson et al., *Social Biases in NLP Models as Barriers for Persons with Disabilities*, arXiv (2020), available at https://arxiv.org/abs/2005.00813
    - *How can you design for accessibility?*
        - B. Oyewole, *Making Accessibility the Default*, Medium (2018), available at https://medium.com/@busayomi/making-accessibility-the-default-f8a9f1d8e259.

        - K. White, S. Abov-Zahra, and S. Lawton Henry, *Designing for Web Accessibility*, Web Accessibility Initiative (2019), available at https://www.w3.org/WAI/tips/designing/.
        - N. Daniels, *Lesson of the Day: 'The Hidden Image Descriptions Making the Internet Accessible,'* The New York Times (2022), available at https://www.nytimes.com/2022/02/28/learning/lesson-plans/lesson-of-the-day-the-hidden-image-descriptions-making-the-internet-accessible.html.
    - *Disabilities and data privacy*
        - K. Hao, *Can you make an AI that isn't abelist?*, MIT Technology Review (2018) **[NB: This article will be made available through LEARN]**
    - *Mental health and data privacy*
        - A. S. Levine, Suicide hotline shares data with for-profit spinoff raising ethical questions, Politico (2022), available at https://www.politico.com/news/2022/01/28/suicide-hotline-silicon-valley-privacy-debates-00002617.
        - T. Germain, *Mental Health Apps Aren't All As Private As You May Think*, Consumer Reports (2021), available at https://www.consumerreports.org/health-privacy/mental-health-apps-and-user-privacy-a7415198244/.

9. Tuesday, March 14: Global Issues Including Government Surveillance

- *Introduction to Surveillance Capitalism*
    - *In a nutshell: Shoshana Zuboff – Surveillance Capitalism and Democracy*, Alexander von Humboldt Institut for Internet and Society, YouTube (2021), available at https://www.youtube.com/watch?v=5AvtUrHxg8A.
    - J. Laidler, *High tech is watching you*, The Harvard Gazette (Mar. 4, 2019), available at https://news.harvard.edu/gazette/story/2019/03/harvard-professor-says-surveillance-capitalism-is-undermining-democracy/.
- *Policing and the police state*
    - R. Rajagopalan, *This is What a 21st-Century Police State Really Looks Like*, BuzzFeed News (2017), available at https://www.buzzfeednews.com/article/meghara/the-police-state-of-the-future-is-already-here.
    - X. Wang, *Hundreds of Chinese citizens told me what they thought about the controversial credit score system*, The Conversation (2019), available at https://theconversation.com/hundreds-of-chinese-citizens-told-me-what-they-thought-about-the-controversial-social-credit-system-127467.
    - N. Köksal, *'Terrifying': How a single line of computer code put thousands of innocent Turks in jail*, CBC (2018), available at https://www.cbc.ca/news/world/terrifying-how-a-single-line-of-computer-code-put-thousands-of-innocent-turks-in-jail-1.4495021
- *Data colonization*
    - E. Erastus, *Algorithmic Apartheid? African Lives Matter in Responsible AI Discourse*, Paradigm Initiative (2021), available at

https://paradigmhq.org/algorithmic-apartheid-african-lives-matter-in-responsible-ai-discourse/).
    - A. Hawkins, *Beijing's Big Brother Tech Needs African Faces*, Foreign Policy (2018), available at https://foreignpolicy.com/2018/07/24/beijings-big-brother-tech-needs-african-faces/
- *Choose your own adventure*
    - Go to the following website and poke around the examples of corporate and government data abuse:
      Privacy International Abuse Tracker, available at https://privacyinternational.org/examples.

10. Tuesday, March 21: Potential Fixes – Greater Inclusivity in the Tech Sector and in CS Education – Can We Change Culture?

- *Perceptions and self-perceptions in STEM*
    - B. Bloodhart et al, *Outperforming yet undervalued: Undergraduate women in STEM*, PLOS ONE (2020), available at https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0234685.
    - M. Campbell, *Feeling Like a Fraud: Imposter Syndrome in STEM*, Technology Networks (2019), available at https://www.technologynetworks.com/tn/articles/feeling-like-a-fraud-impostor-syndrome-in-stem-324839
- *Experiences in the workplace*
    - J. Bhuiyan et al., *Black and brown tech workers share their experiences of racism on the job*, Los Angeles Times (2020), available at https://www.latimes.com/business/technology/story/2020-06-24/diversity-in-tech-tech-workers-tell-their-story.
    - S. Fowler, Reflecting on One Very, Very Strange Year at Uber, Blog (2017), available at https://www.susanjfowler.com/blog/2017/2/19/reflecting-on-one-very-strange-year-at-uber.
- *Changing culture in CS education*
    - J. Dai, *The Paradox of Socially Responsible Computing: The limits and potential of teaching tech ethics*, The College Hill Independent (2020), available at https://www.theindy.org/2235.
    - L. Blum and C. Frieze, *The Evolving Culture of Computing: Similarity Is the Difference*, Frontiers (2005), https://www.researchgate.net/publication/249903204_The_Evolving_Culture_of_Computing_Similarity_Is_the_Difference.
    - C. Frieze and J. Quesenberry, *Change culture not curriculum to get more women into computer science*, re:Work (2016), available at https://rework.withgoogle.com/blog/change-culture-not-curriculum-for-women-in-cs/.
      K. Walther & R.E. Ladner, *Broadening Participation by Teaching Accessibility*, Communications of the ACM (2021), available at

https://cacm.acm.org/magazines/2021/10/255707-broadening-participation-by-teaching-accessibility/fulltext.

- *Hiring, retention, and promotion*
  - L. Mundy, *Why is Silicon Valley So Awful to Women?*, The Atlantic (2017), available at https://www.theatlantic.com/magazine/archive/2017/04/why-is-silicon-valley-so-awful-to-women/517788/.
  - I. Bogost, *The Problem With Diversity in Computing*, The Atlantic (2019), available at https://www.theatlantic.com/technology/archive/2019/06/tech-computers-are-bigger-problem-diversity/592456/.

11. **Tuesday, March 28: Potential fixes – Changing Perspectives, Soft Law and Self-Regulation, Hard Law and Government Regulation, and Other Approaches**
    - *Changing perspectives, data, and technology*
      - M. Miceli, J. Posada, and T. Yang, *Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?*, 6 Proceedings of the ACM Human-Computer Interaction 1-14 (2022), available at https://dl.acm.org/doi/abs/10.1145/3492853**.**
      - K. Hao, *The Fight to Reclaim AI*, MIT Technology Review (2021) **[NB: This article will be made available through LEARN]**
      - J. Kleinberg et al., *Algorithms as discrimination detectors*, PNAS (2020), available at https://www.pnas.org/content/pnas/117/48/30096.full.pdf.
      - L.F. Eisenstadt, *#MeToo Bots and the AI Workplace*, University of Pennsylvania Journal of Business Law (2021), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3921186.
    - *Bug Bounties*
      - J. Kenway et al., *Bug Bounties for Algorithmic Harms*, Algorithm Justice League (2022), available at https://drive.google.com/file/d/1f4hVwQNiwp13zy62wUhwIg84lOq0ciG_/view.
    - *Soft law and self-regulation*
      - A. Tardif, *What the White House's AI Bill of Rights Means for America & the Rest of the World*, Unite.AI (2022), available at https://www.unite.ai/what-the-white-houses-ai-bill-of-rights-means-for-america-the-rest-of-the-world/.
      - T. Hagendorff, *The Ethics of AI Ethics: An Evaluation of Guidelines*, Minds and Machines (2020), available at https://link.springer.com/article/10.1007/s11023-020-09517-8.

    - *Hard law and government regulation*
      - N. Thomas, *Regulating AI: Critical Issues and Choices*, Law Commission of Ontario Issues Paper (2021), available at https://bit.ly/3EaRyF3. → **Read only Sections IX and X at pages 17-29.**
      - M. Schaake, The European Commission's Artificial Intelligence Act, Stanford University Human-Centered Artificial Intelligence (2021), available at

https://hai.stanford.edu/sites/default/files/2021-06/HAI_Issue-Brief_The-European-Commissions-Artificial-Intelligence-Act.pdf.

- *Other Approaches*
  - *Algorithmic Impact Assessments and Audits*
    - *Algorithmic Impact Assessment Tool*, Government of Canada (2022), available at https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html.
    - M. Jedreski et al., *New York City's Groundbreaking New Law Will Require Audits of AI and Algorithmic Systems That Drive Employment Decisions*, Artificial Intelligence Law Advisor Blog (2021), available at https://www.dwt.com/blogs/artificial-intelligence-law-advisor/2021/12/nyc-employment-ai-bias-audit-law
  - *Specialized agencies*
    - A. Tutt, *An FDA for Algorithms*, Administrative Law Review (2017), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2747994
  - *Worker power and unionization*
    - K. Hao and N. Freischlad, The gig workers fighting back against the algorithms, MIT Technology Review (2022), available at https://www.technologyreview.com/2022/04/21/1050381/the-gig-workers-fighting-back-against-the-algorithms/

12. Tuesday, April 4: Voluntary Student Presentations of Small-Group Projects and Course Wrap-Up