Economics, Law and Ethics

Part IB CST

2025-26

Lecture 8: Contemporary ethical issues

Alice Hutchings

What ethical issues in computer science do you find concerning, and why?

Overview

- Contemporary ethical issues:
  - Algorithmic decision making and bias
  - Environmental impacts
  - Power of private platforms
  - Gamification and engagement
  - Privacy
  - Ethics in the workplace
  - Ethics of AI business models

Algorithmic decision making and bias

- Algorithmic decision-making:
  - Can be efficient (computer scientists prize efficiency!)
  - Can lead to unwanted effects, e.g. biased decision-making
- Algorithmic bias:
  - Leads to systematic errors that create unfair outcomes
  - Often happens when training data are unrepresentative, incomplete, or otherwise flawed, typically due to historic inequalities.

Algorithmic bias: Justice system

- 2016: COMPAS found to be racially biased
  - Commercial tool used by judges, probation and parole officers to assess the likelihood of re-offending
  - A landmark review found that Black defendants were more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism
- 2025: Automated racism report by Amnesty International
  - Almost three-quarters of UK police forces use predictive policing, e.g. hotspot mapping
  - Concerns about racial profiling, discrimination, over-policing, privacy incursions

Algorithmic bias: Justice system

- White collar offenders and corporate polluters go unpunished, but someone in a ‘bad’ neighbourhood could end up in prison for a minor infraction
- Feedback loops:
  - Imprisonment affects later opportunities (employment, etc.), potentially leading to further interactions with the criminal justice system
  - Over-policing one area creates more crime data → algorithm sends more police → more data

Algorithmic bias: Recruitment tools

- 2018: Amazon scraps recruitment tool after finding it was biased against women
  - Training data were resumes submitted over the previous 10 years
  - Male candidates were preferred, reflecting male dominance in the tech industry
- Bias in LLMs used for resume screening (Wilson & Caliskan, 2024)
  - White-associated names favoured 85% of cases
  - Female-associated names favoured in 11% of cases
  - Black men are disadvantaged in 100% of cases

Algorithmic bias: Healthcare

- Algorithm used widely in the US to guide care decision-making for millions of people (Obermeyer et al., 2019)
  - Black patients with chronic diseases were ranked as equally in need of care as white patients who were less ill
  - The algorithm excluded race as a factor, but the outputs still reflected bias
  - Black patients traditionally face barriers in accessing health care

Algorithmic bias: Social media

- Algorithmic amplification of politics on Twitter (Huszár et al., 2021)
  - Large-scale randomised experiment comparing a reverse-chronological feed with the algorithmically-curated feed
  - In 6 out of 7 countries studied, mainstream political right content was systematically amplified more than the political left

Algorithmic bias: Education

- Exam cancellations during the pandemic led to the 2020 A-level results being decided by algorithms
- Mass downgrading of results
  - Hardest hit were high-achieving students from poorly performing schools
  - Criticised for disadvantaging students from lower socio-economic backgrounds
- Public protests in Scotland and England led to the withdrawal of these grades

ite Downing Street protesting against the downgrading of A-level results o
ɔh: Matthew Chattle/Rex/Shutterstock

Algorithmic bias

- Usually not due to bad intentions
- Imbalanced/unrepresentative training data
- Structure inequalities
- Variables that seem innocuous can act as proxies for race, class, gender or other protected characteristics

Justifying decisions

- Broader problem: many AI systems cannot explain their actions
- Commercial algorithms are proprietary
- “Computer says no” doesn’t work where people have a right to an explanation and many kinds of discrimination are explicitly illegal

Environmental impacts

Environmental impacts

- Tragedy of the Commons
  - Rational for companies to pollute more, as they gain financially, but to the long-term detriment of others
- Technological advances have led to pollution and the depletion of natural resources
  - Industrial revolution → burning of fossil fuels, power stations, mass agriculture & vehicles
  - Air and water pollution: negative impacts on health/climate change
  - Natural resources are consumed faster than they can be replaced (deforestation); mining of fossil fuels/minerals

e-Waste

- Not just production/running, also end-of-life
  - Many devices are used for relatively short periods
  - Companies increase lock-in with fancy new technologies that meet invented needs
- e-Waste is recognised as the fastest-growing waste stream in the industrialised world
  - TVs, phones, computers, peripherals, gaming consoles, printers, cables, batteries, media devices…
  - Toxic substances, hazardous to humans and the environment, e.g. mercury, lead, arsenic, cadmium, beryllium, leach into the ground and eventually the water table

e-Waste

- Solutions start with redesigning products to minimise/eliminate toxic substances
- Valuable raw materials can be recovered
- Some countries (mainly Europe) have legislation extending producers' responsibility for e-Waste recycling
- The Basel Convention bans the export of hazardous waste to emerging economies, but the exports continue

e-Waste recycling

- Displacement to the global south
- Notorious digital dumping grounds expose populations to hazardous waste

Electrical cables and other components are being burnt to melt off the plastic and reclaim copper wiring

A large scrap yard next to a slum in Accra, Ghana, home to around 100,000 people

Power of private platforms

- They shape what captures our attention:
  - Tim Wu's "The Attention Merchants" (2016)
    1830s tabloids → state propaganda → mass media → today's algorithmic platforms monetise attention at all costs
- They shape what divides us:
  - Auction theory tells us about polarisation → social media incentivises clickbait and fake news
  - Social network theory: echo chambers reinforce group views and can accelerate polarisation

Power of private platforms

- They shape politics
  - Platforms are battlegrounds for persuasion, mobilisation, and disinformation
- Incentives are misaligned:
  - Companies profit from sensationalised and provocative content
  - Censoring is unpopular and costs money
- Platforms are not just communication tools, they are political actors

Technology safety concerns

- Safety concerns arising from the use of social media, chatbots, etc., include:
  - Anxiety and depression
  - Sleep disruption
  - Exposure to bullying, grooming, harassment, and harmful content
  - Loneliness and emotional dependence
  - Design features that encourage compulsive use
- Social media companies allegedly demand evidence thresholds for action that cannot be met by the resources allocated to internal teams or external researchers (Orben & Matias, 2025)

Gamification and engagement

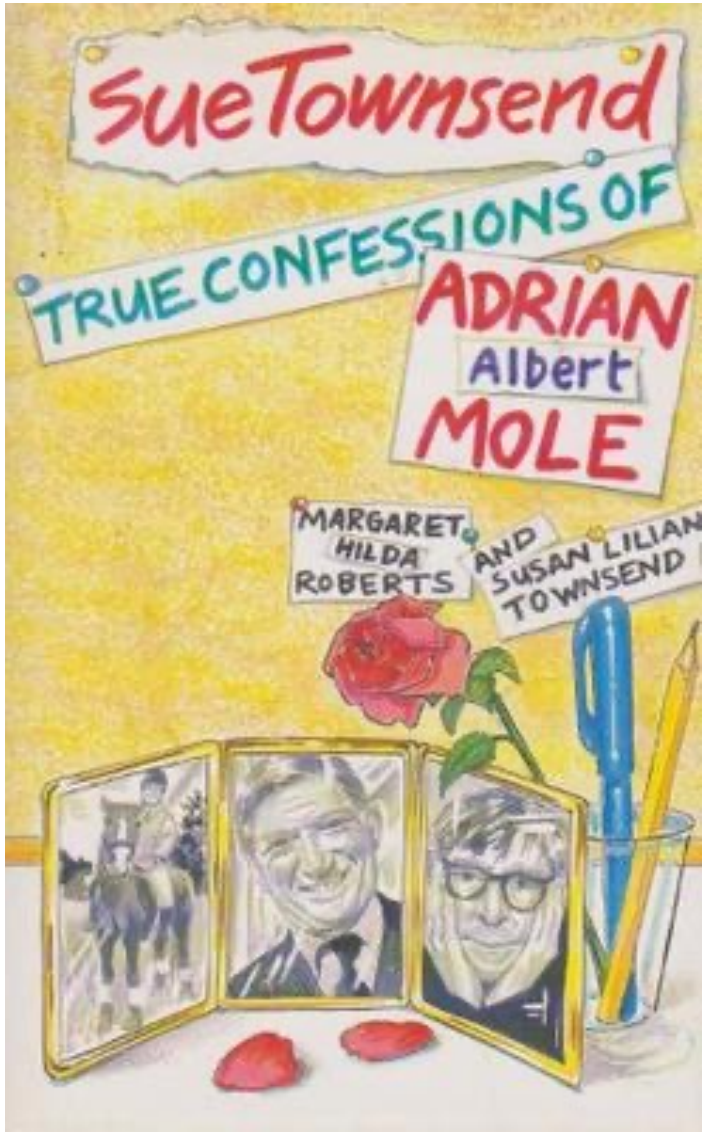
Tim Kendall, Director of Monetization at Facebook (2006-2010) told the US Congress:

"We sought to mine as much attention as humanly possible... We took a page from Big Tobacco's playbook, working to make our offering addictive at the outset. Tobacco companies initially just sought to make nicotine more potent. But eventually that wasn't enough to grow the business as fast as they wanted. And so they added sugar and menthol to cigarettes so you could hold the smoke in your lungs for longer periods. At Facebook, we added status updates, photo tagging, and likes, which made status and reputation primary and laid the groundwork for a teenage mental health crisis"

Gamification and engagement

- User experience is not about ensuring the platform is useful, usable and enjoyable, but about competing for and monetising users
- Information asymmetry: platforms know much more about us than we know about them
- Implications for privacy: the past cannot be forgotten

Adrian Mole



'And please! … please! … send my diaries back. I would hate them to fall into unfriendly, possibly commercial hands. I am afraid of blackmail; as you know my diaries are full of sex and scandal. Please for the sake of our continuing friendship… send my diaries back!'

(Townsend, 1989)

The public v. private face

- We may be revealing more about our private lives to online platforms than we realise (Stephens-Davidowitz, *Everybody Lies: What the Internet can tell us about who we really are,* 2017)
  - Social media: Public, curated view of ourselves; carefully selected photos, no unflattering content
  - Search engines: Personal diary; searches relating to our innermost fears/desires, most fleeting of thoughts, most outrageous ideas, most sensitive information
  - AI chatbots?

Privacy

- Privacy is context-dependent (John et al., 2011)
  - Questionnaires with personal data about drug use, sexual activity, criminal acts, etc. How many questions would participants answer?
  - Random allocation to a website titled ‘How BAD Are U???’ with a cartoon devil logo or ‘Carnegie Mellon University Executive Council Survey on Ethical Behaviors’ with the university crest and detailed privacy assurances

  - Counterintuitively, those in the second treatment group completed less, as privacy had been made salient
  - Those in the first group filled out much more

Privacy

- People often don’t realise how much information they disclose about themselves until there is a data breach
- Data Protection and Freedom of Information laws push back on the flow of data from the weak to the strong
- Privacy is now much harder, with a continuous flow of data (sensors, cameras, microphones, IoT, apps, platforms)

Privacy

- A fork in the road between US and EU laws
- 2014 report on Big Data by the President Obama's Council of Advisers on Science and Technology
  - Can't stop data collectors; can't regulate processors (they claim); so have to regulate uses
  - Problem: US privacy law regulates only a few uses (such as video rentals) and is weak even there

Privacy

- Meanwhile, in Europe…
  - Google v González, European Court of Justice, 13 May 2014
  - Search on Mr González returned a reference to a lawsuit against him in 1998
  - Google said, that's not our problem.
  - ECJ: Oh yes it is! He can ask either the newspaper or Google to take old stuff down

Privacy

- In Europe
  - Privacy is a fundamental human right
  - Laws provide protection from powerful actors (states, companies)
  - Regulation is rights-based, preventative and universal
- In the US
  - No general federal privacy law
  - Privacy is something you ‘agree’ to in T&C
  - By default, companies can collect data unless a specific sectoral law applies (e.g. health, education, financial data…)

Ethics in the workplace

- With the absence of laws and regulators to enforce them, employees have stood up to hold organisations to account
  - Whistleblowers (e.g., Snowden's 2013 leaks about mass surveillance)
  - Walkouts (e.g., at Google in 2018 following alleged sexual harassment and pay inequality)
  - Petitions and demonstrations (e.g., employees at Amazon, Salesforce and Microsoft objecting to their work used for surveillance or separating families at the US border)

Ethics in the workplace

- ‘Tech workers versus Enshittification’ (Doctorow, 2025)
  - Tech workers once had leverage (high demand, mobility), letting them push back against harmful or user-hostile decisions
  - They were the last line of defence against ‘enshittification’, the degradation of tech companies over time, shifting value from users towards platform owners/shareholders
  - Recent mass layoffs have weakened this power, creating a more precarious, replaceable workforce
  - As worker power diminishes, the moral and practical checks provided by labour are disappearing

Existential threat of AI?

- 2023: The Future of Life Institute's open letter
  - Asks for a six-month moratorium on training AI systems more powerful than GPT-4
  - Claims the technology could pose a profound risk to society and humanity, comparable to global pandemics and nuclear war
  - Signed by more than 2,000 people, including Sam Altman, Elon Musk, Kevin Scott, and Geoff Hinton

Existential threat of AI?

- Critical perspective:
  - In 2020, Google fired Timnit Gebru after she refused to withdraw the ‘Stochastic Parrots’ paper, which warned that LLMs can reinforce discrimination against marginalised groups
  - In response to the Future of Life letter, the authors of the Stochastic Parrots argued that focusing on long-term hypothetical risks ignores the actual harms resulting from the deployment of AI systems right now
  - Cognitive bias: hyperbolic discounting

Ethics of AI business models

- Mining personal data (if you're not paying, you are the product)
- EULA contractual coercion: must accept terms to use the service
- Devaluation of creative labour
- Ghost work, low pay and toxic content for Reinforcement Learning from Human Feedback (RLHF) 'Mechanical Turkers'
- Regulatory imperialism/extraction in low-income: imposition of rules, practices and costs on poorer nations in ways that benefit the powerful

Conclusion

- Economic incentives have their limits (misaligned incentives, efficient ≠ just, tragedy of the commons)
- Collingridge dilemma: regulation is hard
- With the rate of technological change, the most pressing social and ethical issues of our time are likely to relate to online platforms and algorithms
- Computer science is slowly beginning to come to grips with the ethical challenges the discipline faces, but there is a long way to go

Thank you ☺