Wk 1

Analyze a Case Study

In the ACM case, a software company (AllTogether) implements a new inline-edit feature (hover shows a pencil icon, etc.) which improves aesthetics and reduces clutter, but has serious accessibility flaws: the controls are hidden except on pointer hover, keyboard navigation fails, screen magnification users struggle, etc. The company has an accessibility policy, a Diversity Advisory Board, risk registers, etc., but leadership is pressured to ship anyway and questions how many blind users really use the product. The feature is released with known accessibility defects; there is some mitigation (documentation, support, rollback option) but still users with disabilities are harmed, and the product discriminates in effect.

Ethical framework analyses

Kantianism / Deontology

Kantian ethics focuses on duties, principles, respect for persons, and acting according to maxims that one can will to be universal laws. Key deontological ideas include not using people merely as a means, respecting autonomy, following moral rules regardless of consequences. Duty to treat all persons with dignity: Software users with disabilities have rights; ignoring their needs or designing in a way that excludes them violates their dignity. Duty to follow rules / policies: The company has internal policies, company rules (accessibility policy, WCAG conformance), and external norms (WCAG guidelines). From a deontological view, adherence to rules is morally mandatory (unless conflicting duties, but here there's no strong conflicting duty like preserving life). Universalization: If every software product ignored accessibility for speed/market pressures, many people would be harmed; that seems not a maxim you'd want to universalize. Thus, under Kantianism, the case suggests the company's decision to release with known accessibility defects is unethical: it violates duty to treat all users fairly, duty to follow agreed standards, and treats disabled users as means (or excluded) rather than ends.

It is quite conclusive: deontology would say the company had obligations (to include accessibility from the start) which they failed to fulfill. Even mitigation after the fact (support, documentation) does not fully satisfy the duty violated by releasing something known to be discriminatory/inaccessible.

Virtue Ethics

Virtue ethics emphasizes character, virtues (honesty, fairness, compassion, justice, duty) rather than rules or outcomes directly. It asks: what would a virtuous agent do? What kind of character is shown by this decision? The virtues relevant include justice, compassion, integrity, prudence (practical wisdom). Releasing a feature while knowing it harms or excludes disabled users indicates a lack of compassion and justice. It suggests a character more concerned with expediency or image than with fairness. On the other hand, the company does show some integrity by having policies, guidelines, acknowledging defects, offering support, etc. But these seem more reactive than proactive, and do not compensate fully for the harm or exclusion. Thus, virtue ethics also leans toward the conclusion that the decision was not ethically praiseworthy. It would consider the action lacking in virtue, or at least flawed. Unlike

deontology, virtue ethics may allow more nuance: maybe some of the people in leadership believed they were doing what was practical (prudence), or that support, disclosure, etc. ameliorates some deficiencies. But overall, a virtuous company would have built accessibility in from the start.
I see this framework as reasonably conclusive in judging the decision unethical (or at least morally deficient).

Utilitarianism

Utilitarianism asks: what action maximizes overall happiness or wellbeing (minimizes overall harm)? Effects, consequences, including both immediate and longer-term, and matters of number, intensity. Positive consequences: shipping on schedule, perhaps market advantages, satisfied customers who value aesthetics and streamlined UX; possibly lower costs short-term. Negative consequences: excluded or harmed users (those with disabilities) who lose productivity, can't use features, maybe disappointed; possible reputational damage; increased support costs; customers rolling back; possibly legal or regulatory risk; long-term loss of trust. If we weigh them, the harms seem serious: for disabled users, the inability to do work, frustration, exclusion etc. And the number of people potentially affected (even if relatively small compared to total users) is nontrivial. Also, the reputational harm, rollback costs, etc. may offset short-term gains.
Thus, a utilitarian analysis likely concludes the decision was not ethical (i.e. negative utility): more harm than benefit. One might argue a version could be constructed where the benefit (speed to market, revenue, etc.) outweighs the harm if the number of disabled users is very small and if mitigation is strong; but given the case as told, mitigation is partial and harms are clear.
Hence, utilitarianism also points toward unethical decision. It suggests the company should have fixed the accessibility issues before the release so as to maximize well-being.

Conclusion

All three frameworks, in this case, seem to point in the same direction: the decision to launch with known accessibility flaws is ethically problematic / morally wrong,
I find Virtue Ethics particularly compelling. Virtue Ethics appeals because it brings in the dimension of moral character: not just "did we follow the rule?" or "did we maximize welfare?" but "what kind of people/organization are we?" I like that idea of integrity, caring, empathy, being forward-looking. In technology, where rules and outcomes are often murky, virtue ethics helps keep focus on values: fairness, inclusion, and responsibility.

Wk 2

Recommender Systems

Recommender systems on platforms such as YouTube raise key ethical concerns when viewed through the lens of virtue ethics, which focuses on the development of moral character and the cultivation of virtues like honesty, fairness, and self-restraint. Unlike approaches that emphasize duties or outcomes, virtue ethics considers whether actions promote good habits and personal growth. In this light, the ethical evaluation of YouTube's recommendations depends not just on how the system functions, but on the kind of behavior it encourages in both users and designers. When algorithms are built to maximize engagement and advertising revenue, they often appeal to habits of impulsivity rather than reflection. This design can encourage overconsumption, reinforce narrow perspectives, and reduce opportunities for thoughtful choice. Political content, for example, may become more extreme over time because emotionally charged or sensational videos tend to attract longer viewing. Although this keeps users on the platform, it can discourage intellectual humility and openness to opposing views. In terms of virtue ethics, such outcomes do not support the development of a fair-minded or civically engaged audience. Instead, they risk fostering traits such as rigidity, confirmation bias, and even division.

Musical recommendations can follow a similar pattern. While personalized playlists can be enjoyable and expose users to new artists or genres, they also present ethical risks if used primarily to deliver more advertisements or extend screen time without regard for users' long-term well-being. A constant stream of suggestions, especially when driven by commercial motives, can diminish self-regulation and lead to passive consumption. In this case, the system may weaken virtues like moderation and autonomy, replacing them with dependency on automated choices.

Nevertheless, virtue ethics does not imply that recommender systems are inherently unethical. If designed with attention to moral development, these tools can help users cultivate curiosity, cultural literacy, and reflective engagement. Features that promote transparency, allow user control, and introduce diverse perspectives can support responsible decision-making and personal growth. For platforms like YouTube, ethical responsibility includes more than preventing harm. It involves actively designing systems that support the flourishing of individuals by reinforcing habits that align with well-formed character. Recommenders that prioritize depth over duration and understanding over mere attention have the potential to encourage not only better content, but also better people.

Right to be forgotten

Yes, there should be a Right to Be Forgotten in my nation of origin and residence. In today's digital age, personal information can remain accessible online for years, even after it becomes outdated, irrelevant, or misleading. This can cause serious harm to a person's reputation, job prospects, and personal relationships. Individuals should have the ability to request the removal of such information when it no longer serves a legitimate public interest. Ethically, this right is supported by deontological principles that emphasize respect for individual dignity and autonomy. People should have control over their personal data, especially when it affects their

ability to move forward in life. The continued online presence of irrelevant or harmful information often results in unjust consequences for individuals, even when the information no longer reflects who they are or the lives they now lead.

At the same time, the implementation of the Right to Be Forgotten should be carefully balanced with other values such as freedom of expression and the public's right to access factual or historical records. While I believe the right should be universally recognized in principle, its application should vary depending on legal, cultural, and political contexts. A universal baseline could establish that everyone has some degree of control over their digital identity, but specific guidelines should be tailored to each region. For example, countries with strong commitments to press freedom might place more weight on public access to information, while others may prioritize personal privacy. The European Union's General Data Protection Regulation provides a useful model by allowing individuals to request the removal of certain personal information from search engine results, but with exceptions in place for public interest, journalism, and legal matters. My position is grounded in both utilitarian and rights-based ethical reasoning. From a utilitarian view, allowing people to remove harmful or irrelevant content can enhance mental well-being and promote fairness in employment and social opportunities. From a rights-based standpoint, individuals have a moral claim to privacy and control over their personal data. With proper oversight and context-sensitive implementation, the Right to Be Forgotten can offer protection without undermining democratic values or public accountability.

Wk 3

2 Articles

Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I. (The New York Times, 2020)

This article examines the controversial firing of Dr. Timnit Gebru, a prominent artificial intelligence (AI) ethics researcher, from Google. Gebru, known for her groundbreaking work on algorithmic bias and fairness, co-led Google's Ethical AI team and was a respected voice advocating for diversity and accountability in tech. The incident began when Gebru co-authored a research paper that raised ethical concerns about large language models—AI systems that require vast computational resources and datasets. Her paper argued that these systems not only consume enormous amounts of energy, contributing to environmental harm, but also risk amplifying racial and gender biases embedded in the data they are trained on. When she questioned Google's decision to block publication of the paper unless she retracted her name, she was dismissed.

What stands out most in this case is how an internal conflict over ethical transparency led to the silencing of a researcher whose work directly addressed the harms that AI could perpetuate. Gebru's firing sparked widespread outrage across the tech and academic communities, as many saw it as a troubling example of corporate interference in scientific inquiry. It revealed the tension between Google's stated commitment to responsible AI and its commercial interest in advancing the same technologies under scrutiny.

The central ethical issue revolves around the clash between corporate priorities and academic freedom. When a corporation funds or employs researchers but restricts what they can publish, it creates a dangerous precedent—one that places profit and reputation above truth and social responsibility. This case highlights how power dynamics within major technology firms can stifle independent research, especially when findings challenge dominant narratives or threaten profitability.

Moreover, the incident underscores broader systemic issues within the tech industry: the marginalization of women and people of color in leadership positions and the lack of transparency in AI development. Gebru's dismissal serves as a cautionary tale about the need for stronger protections for researchers and more robust oversight of corporate influence in AI ethics. Ultimately, the episode calls for an urgent reevaluation of how we define ethical responsibility in technology—and who gets to make that determination.

Uber's Lesson: The Hard Truths About Bad Behavior (The New York Times, 2017)

This opinion piece discusses the toxic culture at Uber under CEO Travis Kalanick, marked by aggression, harassment, and a "win-at-all-costs" mentality. What I found most compelling was how deeply systemic the issues were—more than just one leader's personality, it reflected an

entire company ethos that rewarded bad behavior. It's a clear example of how company culture can shape and even encourage unethical actions, particularly when driven by hyper-competitive, unregulated startup environments.

The central ethical issue here is the moral responsibility of leadership to create a healthy and ethical workplace. Uber's failure to do so not only harmed employees but also the broader public perception of tech companies. It highlights the importance of ethical leadership and the dangers of prioritizing growth over people's well-being.

Wk 4

Facial Recognition

Given the current imperfect but improving state of facial recognition software, where (if at all) do you think it should and/or should not be used?

In its current state, where facial recognition technology still exhibits errors, especially with people of color, women, and other marginalized groups, its use should be extremely limited and highly scrutinized. For example, law enforcement use, such as identifying individuals for questioning or suspected crimes, should be restricted or paused altogether unless absolutely necessary and backed by human verification and legal oversight. The risks of false positives, wrongful arrests, and biased targeting outweigh the benefits, particularly when these systems disproportionately misidentify people from underrepresented backgrounds.
However, more benign applications, like using facial recognition to take attendance in large college lectures, could be acceptable, provided there are clear consent procedures, transparency, and opt-out options. In such low-stakes environments, the risks are relatively minor, though privacy concerns still remain. Even here, institutions should ensure data is securely stored and not repurposed without consent. Overall, its use in high-stakes contexts with real consequences should be avoided until the technology becomes more accurate, unbiased, and accountable.

Suppose facial recognition improved to a point where identifications are nearly flawless. Then, where (if at all) do you think it should and/or should not be used?

If facial recognition reaches a point where identifications are nearly flawless and equitable across demographic groups, then its use could become more ethically justifiable, but not automatically so. For example, it could potentially be used by law enforcement to identify individuals wanted for serious crimes, provided this is done with judicial oversight, clear legal limits, and strong protections against abuse. Similarly, using it to track suspected terrorists in sensitive areas, like airports or public venues, might be ethically defensible if strict safeguards are in place to protect civil liberties and prevent misuse.
That said, just because a technology works well doesn't mean it should be used everywhere. Even a highly accurate system could be problematic in terms of mass surveillance, such as tracking individuals in public spaces without their knowledge or consent. This raises serious concerns about privacy, autonomy, and the potential for government overreach. In more casual or everyday settings, like unlocking phones or managing check-ins for events, facial recognition might be acceptable with consent, but its use should always be balanced against the right to remain anonymous in public. Accuracy alone doesn't solve the ethical questions around power, consent, and civil rights.

Wk 5

Under what conditions do you think it is okay to change the capabilities of a human being via gene editing and/or neurological intervention?

I believe that gene editing and neurological interventions can be ethically acceptable under specific, well-defined conditions—primarily when the purpose is therapeutic rather than enhancement-oriented. When these technologies are used to treat or prevent serious medical conditions such as cystic fibrosis, Huntington's disease, or sickle cell anemia, they serve a restorative function: helping individuals achieve a normal standard of health rather than surpassing it. In these cases, gene editing offers the possibility of reducing immense suffering, improving quality of life, and even saving lives. From an ethical standpoint, such use aligns with the medical principles of beneficence (doing good) and nonmaleficence (avoiding harm), provided that safety, efficacy, and informed consent are ensured.
Neurological interventions, such as deep brain stimulation for Parkinson's disease or depression, can similarly be justified when they aim to restore lost function or alleviate severe, treatment-resistant conditions. These applications honor the autonomy and dignity of patients, giving them a chance to live fuller, more independent lives. However, ethical approval for these interventions depends on rigorous oversight, transparent communication of risks, and respect for the individual's right to choose freely without coercion or social pressure.

The ethical tension becomes more pronounced when gene editing or neural modification shifts from healing to enhancement—for example, attempts to increase intelligence, memory, or physical strength beyond the natural human range. Such uses raise profound questions about fairness, equality, and the essence of human identity. If enhancement technologies were available only to the wealthy, society could become divided into genetic "haves" and "have-nots," amplifying existing inequalities. Moreover, the pursuit of perfection could erode diversity and individuality—qualities that define humanity's richness.
In my view, the ethical boundary lies between healing and redesigning. Medicine should strive to restore and empower, not to create artificial hierarchies of ability or worth. While research into these technologies should continue under strict ethical and regulatory frameworks, their use should always be guided by compassion, justice, and humility. Ultimately, gene editing and neurological interventions are powerful tools—but their moral value depends entirely on whether we use them to heal human suffering or to reshape humanity itself.

What do you think the stance of the data science community should be with respect to doing such work?

The data science community has a deep ethical responsibility when engaging in gene editing or neurotechnology projects. These interventions rely heavily on massive datasets, predictive algorithms, and machine learning models all areas where bias, error, and unintended consequences can easily creep in. Therefore, data scientists should not just be passive contributors to these technologies but should take an active role in ethical oversight, asking hard questions about purpose, consent, fairness, and long-term impact.

I believe data scientists should only work on such projects if they're part of a transparent, interdisciplinary, and ethically governed effort. They must advocate for inclusive datasets, rigorous testing, and public engagement. Moreover, they have an obligation to speak out if a project veers into ethically questionable territory, especially when it risks harming marginalized communities or contributing to social inequality. In essence, data scientists should act not just as engineers of possibility, but as stewards of responsible progress.

Future work

Do the people or organizations that are behind the changes which cause jobs to disappear have an obligation to help the people who lose jobs?

Yes, organizations that introduce technologies or business models that result in job losses do have an ethical obligation to support the workers who are displaced. While innovation and automation are often necessary for progress, companies should not ignore the human cost of those changes. Ethically, it's a matter of corporate responsibility and fairness if a company benefits from labor for years and then replaces that labor with machines or software, it owes something to those workers beyond just a severance check. This could include retraining programs, extended healthcare, or job placement services. A just transition ensures that the pursuit of efficiency or profit doesn't come at the expense of workers' dignity and livelihood.

Will we enter a world where there is sufficient wealth for all but not sufficient work for all – and if so, should society provide a guaranteed basic income for all?

We may indeed be heading toward a world where automation and AI create abundance in wealth but reduce the need for human labor, especially in repetitive or routine jobs. In that case, a universal basic income (UBI) could be an ethically and economically sound solution. It would ensure that everyone has access to basic necessities, regardless of whether traditional work is available. UBI could also give people freedom to pursue education, caregiving, creative work, or entrepreneurship without the pressure of mere survival. While it wouldn't solve every problem, it acknowledges a fundamental truth: in an age where machines can do much of the work, human worth shouldn't be tied solely to employment. Society must adapt by prioritizing human well-being over outdated definitions of productivity.