Beyond Thoughtfulness: A Student Perspective on Ethical Engineering[1]

Anonymous[2]

Abstract

In response to recent failures of the technology industry, computer science departments are integrating ethics trainings into their curricula, equipping undergraduates with an applied ethics framework that enables them to reason carefully about the inescapable ethical choices involved in designing any technical system. This focus on thoughtfulness is a commendable shift from existing professional norms. Thoughtfulness on its own, however, is too narrow a paradigm to effectively address the structural problems surrounding contemporary technology production. In this commentary, I justify this claim by reflecting on my experience as a computer science major at Harvard and my encounters with different paradigms of ethical engineering. First, I describe what I call thoughtful engineering in the CS department, which emphasized anticipating and avoiding unintended consequences on the level of the individual engineer. Second, I lay out engineering as advocacy in the Government department, which focused on how technologists may best utilize their expertise to educate the public. Third, I characterize the approach of the History of Science department as reflexive engineering and refusal, which considered science and engineering as a social practice and challenged technological determinism. Finally, I introduce inclusive engineering, as I witnessed it in student organizations like Harvard Women in Computer Science, which aimed to broaden the idea of who can be an engineer. I argue that CS departments could benefit from incorporating these additional perspectives into their curricula. In particular, I find the analysis of power relations—as they exist in society and as technical systems may reconfigure them—to be indispensable. Such an analysis undoubtedly requires a deeper engagement with politics than mere thoughtfulness and may therefore best be implemented by CS departments in close collaboration with student-led organizations and partner departments such as Government, STS, and History of Science.

---

[1] This article was accepted as part of a *Journal of Social Computing* special issue on "Technology Ethics in Action: Critical and Interdisciplinary Perspectives." However, it had to be withdrawn because the journal could not publish the article with an anonymous author.

[2] The author is a recent graduate of Harvard College who majored in Computer Science and who now works at a major Silicon Valley technology company.

Keywords: technology ethics; ethical engineering; applied ethics; computer science pedagogy; government; STS; history of science

1 Introduction

Technology ethics is en vogue today. Industry groups, consumer protection agencies, and national governments alike have published AI ethics guidelines and principles in recent years. Several research institutes for AI ethics have been founded at prestigious Western universities[1][2][3][4]. Some observers have even (ironically) dubbed ethics "the hottest product in Silicon Valley"[5]. But what exactly does it mean for practitioners—those who go by titles such as "software engineer" or "data scientist"—to act ethically or to practice ethical engineering?

This is a question I began asking myself as a Computer Science major at Harvard College, and have continued to consider through my subsequent transition into professional life working in the technology industry. My interest in the social impacts of computing led me on a journey that spanned different departments, schools, and student organizations across Harvard. In each of these spaces, I observed different paradigms of ethical engineering. Some programs in the college's Computer Science department, for instance, focused on equipping soon-to-be software engineers with a framework for ethical reasoning. In collaboration with the Philosophy department, the curriculum helped students identify and think through ethical challenges they would encounter in their everyday work—an approach that I call thoughtful engineering. Some professors in the Government department encouraged their students to take an active role to ensure that technologies are developed in the public interest by employing their technical skills to hold the technology industry accountable and by making their expertise available to civic society—which I refer to as engineering as advocacy. Classes in the History of Science department helped students think more theoretically about the ways in which society and technology shape one another, and whether the development of certain technologies is inevitable—leading to what I call reflexive engineering and the possibility of refusal. Lastly, student organizations such as Women in Computer Science advocated for a more inclusive engineering culture that is open to historically marginalized groups.

As I explored these different discussions of tech ethics, I found that their respective conceptions of ethics were seldom brought into conversation with each other. In some places, such as the CS department, ethical engineering was considered to first and foremost be a set of values and practices that individual practitioners could subscribe to. Others, most notably the History of Science and Government departments, took a broader view of technology production, situating ethical engineering in a structural and collective context.

In this article, I will describe four different paradigms around ethical engineering[3] that I observed at Harvard: thoughtful engineering, advocacy, reflexive engineering and refusal of engineering, and inclusive engineering. I will describe the environment in which I came in contact with each of them, outline what I perceived to be each environment's notion of ethical engineering, and some of the limitations that I noticed about each of these paradigms.

For me, learning to approach these questions of ethical engineering through many different lenses was one of the most exciting and transformative intellectual experiences that I had as an undergraduate—so much so that I have not stopped thinking about them, and am now writing this article to reflect on the different programs' frames of reference. My exploration made me wonder what kinds of individual behaviors I might personally want to adopt as a software engineer and what kinds of institutional arrangements and processes could lead to forms of technology production that better serve the public good. Each conception of tech ethics came with notable insights as well as limitations in this respect. My aim is not to attack these programs (nor the people involved in their creation) by pointing out what I perceive to be their particular shortcomings, but rather to highlight the need to consider tech ethics from a variety of perspectives and the potential of bringing these different viewpoints together towards a more comprehensive discourse of ethical engineering.

[3] I will refer to practitioners using the term "engineers" in this article, but want to encourage readers to interpret it in the broadest sense, encompassing a whole array of people involved in the creation of technical systems, including but not limited to software engineers and data scientists and irrespective of higher education degrees or specific fields of study.

2 Computer Science Department: Ethical Engineering as Thoughtful Engineering

The most prevalent understanding of ethical engineering that I encountered in the computer science department was that of ethical engineering as thoughtful engineering. According to the thoughtfulness narrative, software engineers have historically built and deployed technologies with little consideration of their adverse effects, failing to anticipate potential malicious uses or unintended consequences of their work. Inadvertently, they had thereby contributed to various societal harms such as discriminatory treatment of minorities[6][7][8][9] or foreign election interference[10][11][12]. If only technologists had spent more time thinking rigorously about the downstream effects of their work (beyond making wildly optimistic prognostications), the thinking goes, they could have prevented bad outcomes. As technology pervades ever-larger realms of our daily lives, this kind of negligence is likely to lead to ever-greater harms. Consequently, the urgent task of current-day computer science programs is to make engineers more aware of the stakes of their work and to instill in them a better understanding of their responsibility.

I experienced the CS department's implementation of the thoughtfulness approach in two primary ways: through ethics modules that had been integrated into existing CS classes, in a collaboration between CS faculty and philosophy PhD students, and through newly created classes dedicated entirely to issues surrounding ethical engineering. While the latter approach allowed already interested students to deepen their understanding of ethical engineering, the former assured that virtually everyone who went through Harvard's CS program would be exposed to at least some ethics training, in line with the department's stated goal of equipping students with the necessary skills to "solve problems cooperatively and in an ethically principled way"[13].

As part of Harvard's Embedded EthiCS initiative, philosophy PhD students would deliver guest lectures in a range of CS classes to help students think normatively about their responsibilities[14]. In CS 136, Economics and Computing, for instance, we took one class meeting of the semester to step back from calculating equilibria in zero-sum games and understanding distributed ledgers, and instead talk about the moral obligations we may have when designing recommender systems. A well-known example of a recommender system is the algorithm that YouTube uses to suggest the next video you might be interested in watching. These recommendations determined 70% of the one billion hours of video that people watched on the platform every day in 2017[15]. Recommender systems such as that of YouTube have recently been alleged to contribute to political radicalization and to spread dangerous and inaccurate health information about vaccines[16][17][18][19][20].

In CS 181, Introduction to Machine Learning[21], another course that integrated an ethics module, we took a break from calculating multivariate derivatives and understanding gradient descent to instead think about the fairness of a hypothetical hiring algorithm that evaluated job applicants on the basis of historical performance data. In our particular case study, the dataset suggested that “African-American sales representatives” have historically had “significantly fewer average sales than white sales representatives,” leading the algorithm to recommend hiring disproportionately “fewer African-Americans than white applicants”[22]. The guest lecturers would guide us through these case studies and facilitate a discussion about the ethical choices involved. In the case of the hiring algorithm, we might, for instance, choose to formalize our desire for fairness via a “non-discrimination criterion”[23] to use as a constraint in optimizing the hiring algorithm. Yet even this choice involved further ethical consideration: we could implement non-discrimination via demographic parity (“the decision should be independent of the protected attribute”), equalized odds (“the prediction and attribute should be independent, conditional on the outcome”), or a well-calibrated classifier (“the outcome and protected attribute are independent, conditional on the predictor”).

As we were debating which of these criteria (if any) is just, the realms of probability theory and moral philosophy blurred in a way that felt quite disorienting to me, given how compartmentalized I experienced these two academic disciplines most of the time. Our conversations in class certainly dispelled, in my eyes, the notion that technology is, or ever can be, neutral, and therefore challenged the idea that practitioners could evade responsibility by somehow not putting our “thumb on the scales.”

The ethics units would also at times unearth epistemic differences between the humanities and STEM disciplines. Weighing competing notions of preference-utilitarianism and freedom of choice was unfamiliar territory for many of those who majored in Computer Science, Applied Math, or Statistics. In class discussions, students often were visibly uncomfortable about being asked to reason with definitions and concepts that were perceived to be “vague” or “arbitrary”, especially in contrast with the mathematical ones we were used to encountering. “Who gets to decide what counts as ...?”, “Where should we draw the line?” were questions that were often uttered. Given the (relative) lack of easily operationalizable terminology, as time went on our conversations would frequently not converge on definite agreeable answers, but rather surface even more questions.

This is not to say that looking at a problem through different lenses is a mode of inquiry that is foreign to computer science as a discipline. To the contrary, many important theoretical results in the field were achieved by means of mathematical reductions, i.e. by rephrasing a known problem in terms of another well-understood problem which then would allow for an elegant solution[24]. The difference, that I perceived, for instance in our discussion with ethicist Danielle Allen[4] about recidivism algorithms in the criminal justice system, was that additional angles—for example examining the history of mass incarceration in the United States, and the construction of Black criminality[25]—would make the problem harder to solve.

In addition to the lectures, accompanying problem sets and writing assignments prompted us to think more about unforeseen consequences in order to prevent harm, a practice at least partially at odds with prevalent tech industry mantras such as “move fast and break things.” Overall, I believe these modules, complementing the dedicated semester-long tech ethics classes, succeeded in providing students with frameworks to think normatively about our work as thoughtful practitioners.

Thoughtfulness on its own, however, is insufficient to actually effect positive change, as many students would notice during their summers working as interns in the tech industry. Removed from the classroom case studies and thought experiments that revolved around an individual engineer, we now found ourselves as members of small teams that were part of larger organizations that were part of broader global industries and markets. What am I supposed to do if I find that a certain configuration of my company’s recommender system may unfairly disadvantage one group? What if that system is not part of my specific team’s codebase? Given that my colleagues all have their own perspectives

[4] This discussion was part of a semester-long course focused on critical thinking in data science, rather than an EthiCS module.

and political beliefs, am I merely imposing my own politics onto them? How should I talk to my manager to raise my concerns—will this jeopardize my next performance evaluation? My co-worker is on a company-sponsored H-1B visa. Can she weigh in on a controversial issue, given the risk of her being fired and required to leave the country? Would my concern simply be ignored if remedying the issue would make the product less profitable? None of these questions were addressed by the CS department's ethics units which, by their particular mode of engagement with ethics as moral philosophy, had abstracted away the structural context in which contemporary technology is produced. The thoughtful engineering approach on its own, while undoubtedly broader in scope than any traditional strictly "technical" CS curriculum, was still too narrow to equip us with strategies to hold the organizations that we work for accountable or to effect change within those organizations.

A similar "emphasis on personal responsibility" has been found in the language of professional ethics codes for the responsible development of artificial intelligence. By "position[ing] the individual practitioner as the primary locus of ethical responsibility" ethics codes and the CS department's thoughtful engineering approach alike ask the individual engineer "to take the brunt of ethical conflict and adjudicate between potentially conflicting or incommensurate values" while simultaneously providing "no guidance on how to navigate moral conflicts"[26]. This absence of guidance highlights a gap in the theory of change that underlies the thoughtfulness approach. In the prescient words of Langdon Winner: "According to a very common and laudable view, part of the education of persons learning advanced scientific skills ought to be a full comprehension of the social implications of their work. Enlightened professionals should have a solid grasp of ethics relevant to their activities. But, one can ask, what good will it do to nourish this moral sensibility and then place the individual in an organizational situation that mocks the very idea of responsible conduct?"[27] How are the thoughtful technologists that come out of tech ethics programs supposed to actually create better, more ethical technology?

3 Government Department: Ethical Engineering as Advocacy

Another dimension of ethical engineering that I encountered, predominantly in the Government department, also centered harm-reduction as a core principle, but differed from the thoughtfulness approach in terms of the methods it deployed in pursuit of that goal. Through its Tech Science program, Harvard's Government department encouraged soon-to-be technologists to practice ethical engineering as advocacy[28]. Computer scientists should not limit their concerns to making thoughtful choices in the development of whichever system they have been tasked with building, but, in addition, should conceive of themselves as subject matter experts who are uniquely positioned to educate the public about the capabilities and limitations of digital technologies. Empowered by such expert advice, the public at large can then make informed choices about how it wants (or does not want) to deploy certain technologies. Putting the onus of explanation on the technologists themselves seemed logical to me. If advanced technologies are, as one lecturer suggested, indeed, per Arthur C. Clarke, "indistinguishable from magic," who else but the magicians themselves have the ability, and thereby the obligation, to explain the tricks?

There is a rich history of computer scientists using their expertise to raise concerns about the use of technology, practicing "computing as rebuttal"[29], perhaps most prevalent in the field of computer security. In the past, for example, those practicing ethical engineering as advocacy have argued against the use of electronic voting machines[30][31], resisted governments calling for the impositions of backdoors into encrypted systems[32], and challenged the practice of governments stockpiling software vulnerabilities[33]. More recently, computer scientists have joined multidisciplinary teams with, for example, legal scholars to highlight civil liberties concerns around the use of facial recognition systems by law enforcement agencies[34].

Engineering as advocacy is qualitatively different from "traditional" software engineering. Rather than predominantly building and maintaining large software systems, engineer advocates may craft bespoke demonstrations that illustrate matters of public concern. In 1997, for instance, Latanya Sweeney famously established that only three pieces of information—date of birth, gender, and ZIP code—were sufficient to uniquely identify 87% of U.S. citizens[35]. To illustrate the importance and implications of her finding, Sweeney, then a graduate student at MIT, re-identified the medical record of then

Governor of Massachusetts, William Weld, from supposedly "anonymized" medical data. She did so by cross-referencing "quasi-identifiers" via publicly available voter records. In another instance of ethical engineering as advocacy in 2006, members of the German Chaos Computer Club (CCC) and Dutch security researchers jointly demonstrated the ease with which electronic voting machines could be compromised to manipulate vote counts by probing a device they had managed to acquire from a voting machine manufacturer[36]. Even more recently, in 2018, Joy Buolamwini and Timnit Gebru showed how predictions of commercial facial recognition software were systematically less accurate for people with darker skin tone than for those with lighter skin tone[37].

By crafting a straightforward narrative around their research, and effectively communicating it, technologists who practice ethical engineering as advocacy can achieve societal impact. Sweeney, for instance, was asked to testify in front of Congress and contributed to the robust privacy protections that were eventually enshrined in HIPAA[38]. Following CCC's advisory, the German federal constitutional court declared the use of voting computers to be unconstitutional[39], and in response to Boulamwini and Gebru's critiques, IBM and Microsoft, among others, made efforts to improve the accuracy[5] of their facial recognition systems[40].

In Professor Sweeney's class at Harvard, students of ethical engineering as advocacy would learn how to devise such simple experiments and how to effectively convey their implications. We would work in small groups to come up with a research question, conduct experiments, and eventually produce a writeup thereof on a weekly basis. What surprised me throughout these quick-turnaround cycles, was how few resources were typically required to conduct such experiments. A single undergraduate could build a browser extension that showed users how much of their private information the payment service Venmo needlessly broadcasts to the public[41] or demonstrate how easily bad actors could use modern natural language processing models to distort civic discourse, flooding federal public comment pages online with a thousand fake comments[42]. In addition to the low-resource requirements, I discovered that simplicity is often a virtue for such

[5] Many have since argued that making vulnerable populations more legible to surveillance infrastructures does harm and called for abolition of facial recognition. See the later sections on refusal and community-led technology design.

demonstrations, as it makes it easier to craft a straight-forward narrative around a concern and about harms that even relatively unsophisticated actors could cause.

Writing these short papers and presenting them in front of our peers in quick iterative fashion helped us students to become better scientific communicators and to proactively structure our research goals and methodology around the public conversation we wanted to spark. Of course, doing this research, often using datasets that reflected people's personal information, came with its own ethical questions—some of which were addressed by a human-subject research training that all of the students in the class had to complete in the first weeks of the semester[43].[6] In line with the Belmont Report[44], which was required reading, we would, for instance, always consider whether our work benefitted the research subjects.

Another welcome side-effect of this flavor of ethical engineering, as I understood it, was that it de-centered the computer scientist in debates about the appropriate use of computing in society. Rather than being the sole arbiter of such debates, a computer scientist engaged in ethical engineering as advocacy acts more as a facilitator who enables people from a diverse set of backgrounds to weigh in and co-design technology. The engineer advocate therefore recognizes that they are not positioned to make decisions about the ethics and equities of technologies they build, but rather that their role is "helping the helpers"[45], for example providing expertise to journalists, public office holders, and regulators.

In software development, people generally try to avoid "reinventing the wheel," by which they mean re-implementing a functionality that already exists. Instead, as a best practice, software engineers re-use functionality provided by standard libraries, often made available via open-source software. This way, they can focus their energy on building the new custom functionality that is specific to the problem they are trying to solve, and otherwise use off-the-shelf components which a whole community might already have spent years creating, optimizing, and making reliable. Similarly, engineers as advocates do not have to start from scratch when imagining how decision-making processes for building software could be made more accountable and legitimate. Instead, they can lean on century-old innovations that have created our public sphere, democratic institutions with elected

[6] The same certification had been required by the CS department class Critical Thinking in Data Science.

representatives, human rights standards, the rule of law, and popular movements to help create broadly legitimate technological futures that serve all of us. This approach was emphasized concretely in the course, where we would simulate congressional hearings, alternating between roles of elected officials, representatives of civil society, industry, and engineer advocates making their case.

Nonetheless, there are some limitations to the advocacy paradigm. For instance, while some of the demonstrations that I listed may require comparatively few resources in terms of compute and personal time and are therefore accessible, there are significant legal barriers to engineering as advocacy. For instance, the U.S. Computer Fraud and Abuse Act may make it illegal to scrape user-generated data from some online platforms[46]. Moreover, instead of valuing the contribution of re-identification research to public security, government entities such as the Australian Department of Health have pressured universities to terminate professorships[47], and even attempted to criminalize re-identification research altogether[48]. Large research institutions such as Harvard could, if necessary, protect their faculty and students when they conduct web scraping for research purposes. Independent researchers who act without such institutional support would not have similar protections available to them and may therefore be precluded from working in some sensitive areas around privacy and security, such as auditing online platforms.

Another pitfall with ethical engineering as advocacy is that, while it may be effective at alerting the public about a given technological issue, it may not be able to do much more than that.[7] Auditing YouTube's recommender system for alleged tendencies to serve its viewers ever-more radical content over time, for instance, there is only so much that researchers can learn working outside of data-rich organizations rather than within them. Advocates may call foul, get the media to pay attention to an issue, and pressure tech companies to investigate a matter internally—which would likely be powered by more resources and more granular data—but in the end, academic researchers remain on the sidelines. Some initiatives such as Harvard's Social Science One[49] aim to alleviate this issue by building partnerships that would allow academic researchers to conduct independent research on data gathered by the tech industry. Nonetheless, such relationships between academia and industry will in all likelihood be fragile, given the different interests

[7] Of course, oftentimes naming a problem is a large contribution to solving it.

of the parties involved, such as conducting original and critical research, complying with privacy regulations, and protecting business reputations.

In addition to legal barriers and lack of available industry data, another challenge with ethical engineering as advocacy is that there are few well-defined career paths available to students. While the transition to working at a large tech company after graduating college was well-understood (and commonplace) among my peers, pursuing ethical engineering as advocacy full-time seemed like a much less certain choice. There appeared to be only a handful of jobs that allow people to work as engineer advocates, for instance in academia, for nonprofits, or as staffers for elected officials[50]. As cryptographer Bruce Schneier writes, some public-interest technologists—which I believe to include engineer advocates— "do this full time as a career. Others take short leaves of absence from their careers to pursue public-interest technology. Still others do this in their spare time, as an avocation"[51]. Hence, building institutional support for ethical engineering as advocacy will be required to realize its full potential.

4 History of Science Department: Ethical Engineering as Reflexing Engineering and Refusal

In the History of Science Department, we looked at science and technology in a way that was very different from that of my STEM classes: as a social practice. Instead of conceiving of scientific and technological history as an accumulation of great ideas and discoveries, history of science put scientific communities themselves under the microscope. While in hindsight this perspective turned out to be quite intuitive and clearly useful, I had never been exposed to it prior to stumbling into an introductory class. This gave me the opportunity to immerse myself in the "social, cultural, political, and economic context in which scientific knowledge has been produced and applied"[52]. The class covered a variety of themes such as the relationship of science and religion illustrated by Galileo's confrontation with the church; efforts to classify the natural world, on the basis of the works of Carl Linnaeus and Thomas Jefferson (the latter having infamously justified white supremacy by categorizing human beings[53]) and the relationship between science and the state during the Cold War.

What kept surprising me throughout the semester was how consistently the scientific practices and understandings that we examined were historically and socially contingent. The standpoint of the people who did the science mattered[54][55]. It influenced the questions they asked and the knowledge they produced—a statement so striking in its simplicity and obviousness that one risks forgetting to state it (and seldom, if ever, hears it in STEM classes). This clearly contrasted with the notion of technological determinism—a term which I also encountered for the first time in History of Science—the theory that technology is an unstoppable force that structures society and that society itself has no capacity to shape. The fact that I didn't even have a name for it prior to taking the course goes to show just how much of a "self-evident feature of modern life"[56] technological determinism is considered to be.

Given this perspective on the past and the theoretical vocabulary to make sense of it, I began to see the present much differently. Scientific "progress" and technological development, in my mind, no longer amounted to a universal, unidirectional, deterministic chain of events[57]. Instead, there appeared to be vast spaces of agency behind any scientific or technological venture, choices that scientists and engineers today make and ought to reflect about. "Reflexive engineers," as Peter T. Robbins calls them, recognize this potential. "Challenged and interested to think about their place in the world a little more clearly,"[58] as an interviewee of Robbins put it, reflexive engineers might ask themselves: Which values are encoded in the way computer science makes sense of the world and are they in alignment with ours? What is our "professional ideology"[58]? What is our "vision of how technology has a place within broader concerns"[58]?

In recognition of this value-ladenness of computer science, one of the possible choices that the history of science brought to my attention led to perhaps the most radical interpretation of ethical engineering that I had come across: refusal to participate in certain scientific or technological endeavors altogether. Having thought through the possible unintended consequences of building a system, a reflexive engineer may decide that, rather than thoughtfully approaching a solution, it may be better not to build such a system at all—thereby becoming a conscientious objector. The consequence of collective action of this kind would be that some technologies would not be built.

In class, I encountered specific accounts of refusal among Cold War scientists and technologists. Their stories in turn reminded me of that of a relative of mine. A new grad software engineer in the 1980s, he worked on a PEARL real-time system for a microcontroller. The main customer of said system, he soon found out, was using it to guide torpedoes deployed in the Falklands War[59]. After voicing his discomfort with colleagues, and learning that management did not see a problem in military contracts and would take on future ones, he eventually quit his job. Demand for software engineers was so high, he recounted, that he could afford to take a stance. While I had heard his story in the past, the perspective that history of science had given me allowed me to understand his actions in a broader context. In similar vein, employees at large tech companies have also lately quit their jobs in protest, one Facebook data scientist stating in a matter-of-fact way in the customary "badge post" upon his departure: "Unfortunately, I don't feel I can stay on in good conscience. (1) I think Facebook is probably having a net negative influence on politics in Western countries […] (2) I don't think that leadership is involved in a good-faith effort to fix this […] (3) I don't think I can substantially improve things by staying"[60].

Such a narrative of a "great refusal" reminded me of the student movements of the New Left in the 1960s and their choice to not contribute to military-related research. As Mario Savio, a student at UC Berkeley, put it back then, "there's a time when the operation of the machine becomes so odious, makes you so sick at heart, that you can't take part, you can't even tacitly take part. And you've got to put your bodies upon the gears and upon the wheels, upon the levers, upon all the apparatus, and you've got to make it stop. And you've got to indicate to the people who run it, to the people who own it, that unless you're free, the machine will be prevented from working at all"[61].[8]

Today, in contrast, in the computer science department, technological determinism was the norm, so much so that it was never even explicitly articulated, let alone challenged, in the courses I took. Conversations about refusal were absent from the classroom where, at most, the possibility of ethical technologists reducing the harm of to-be-deployed technologies was considered. One instance of this occurred in a CS class focused on critical thinking in data science. Students in the course were discussing a case study of a "resume

---

[8] As Turner argues, in the context of the speech, the metaphor of the "machine" is to be understood much more broadly as the role of the university as a provider of professional training for future knowledge workers.

vetting system." In a hypothetical scenario, software developers at a firm had decided to build an automated screening system to reduce the workload of their HR department, which was overburdened with job applications[62]. While the system successfully alleviated the burden on the HR department by filtering the applications to a manageable number of candidates, it also seemed to discriminate against people with disabilities. In addition, the system employed a black-box model and was therefore neither able to explain its decisions nor to be held accountable for them.

In the classroom, students faced with this conundrum spent a long time discussing how these bugs may be fixed through imposing fairness constraints, "de-biasing" the underlying training data, and many other means. The idea of simply hiring an additional HR person to vet incoming applications, rather than spending many hours of engineering time on algorithmic tweaking with questionable prospects of success, was not brought up until the end of the discussion. In fact, when one student raised the question of whether this problem may be better addressed simply through additional staff, the professor responded that such a consideration was out of scope for the course.

Why was refusal out of scope in this situation? The most convincing explanation I could come up with is that conversations about the usefulness of technology are, in a sense, antithetical to the usual mode of thinking in computer science. At the risk of becoming too schematic, "traditional engineers," as Robbins calls them, love to find creative ways to push the envelope in order to solve a problem subject to a set of constraints. Questioning the constraints themselves (e.g., the idea that no additional HR personnel could be hired) is usually not an option—if it were, there might be no engineering work left to do at the end of the day. This left me wondering: how critical can critical thinking in data science truly be if it presupposes the use of technology as the correct answer to any questions?

Environmental-engineer-turned-artist Tega Brain shares this frustration "with the way that engineering as a discipline tends to frame problems as technical challenges. You're supposed to scope out the political and social forces that are causing an environmental problem, and just slap a technical fix on the end of it. [...] It seemed like my job was to make these wildly unsustainable projects just a little less bad"[63]. How can computer science move beyond the attitude that she, in her art, pointedly mocks as "you want us to machine learn our way out of the climate crisis? Lol, let's do it."?

The perspective that history of science has given me on computing made me wonder what it would look like to incorporate that mode of thinking into CS courses, akin to the Embedded EthiCS partnership with philosophy. Unlike the somewhat abstract insights provided by the philosophy department, I suspect (and would hope) that history of science or STS collaborations would require reflexivity on the part of the computer science department itself: asking tough questions about its own complicity—in datafication, administrative violence[26][64], and global inequality accelerated by tech-enabled winner-take-all economies[65], just to name a few—in light of its strong ties to industry and its usually unquestioned embrace of techno-deterministic and techno-solutionist thinking.

There are some signs, at least on some U.S. college campuses, that student sentiment is shifting and that refusal has been reintroduced into the realm of imaginable actions. At Stanford, for instance, the undergraduate collective Students for the Liberation of All People (SLAP) has disrupted career fairs, openly challenging their classmates to reflect on whether they want to work for companies such as Microsoft, Amazon, Salesforce, and Palantir who have direct contracts with US Immigration and Customs Enforcement (ICE)[66]. Students moreover seem to increasingly realize that working for large tech companies may not be all that different from working in investment banking or management consulting—two other high-prestige fields Harvard graduates are drawn to each year[9]—than they had previously believed.[10] However, I personally did not witness either of these views (refusal as a possibility or seeing big tech as highly problematic) to be widespread in Harvard's CS community, where students who announced that they had secured a coveted internship at Facebook or Palantir over dinner would still be met with congratulations and admiration.

What could be done to further normalize refusal as part of ethical engineering? One approach, suggested in the literature, would be to contextualize refusal as a joyful and generative act[69]. After graduation, I personally witnessed an example of this. The team I had joined, following thorough field research and deliberation, decided against launching a feature we had built, as we feared it would cause more harm than good. Rather than

[9] "Of respondents who plan to enter the workforce, a plurality — 18 percent — indicated they are going into consulting. Finance and technology clocked in at a close second and third in terms of popularity, drawing 16 and 14 percent of graduating seniors, respectively."[67]
[10] "Investment Banking, but Worse"[68]

considering this a failure, the project lead announced the no-launch decision in a memo in the same manner in which they would usually highlight a product launch, thanking all those involved in arriving at the decision. Due to my prior exposure to refusal, I made sure to be among those loudly and openly celebrating this occasion, in the hope of contributing my part to a culture in which people acting as "citizen engineers and citizen professionals"[69] are valued.

5 Student Organizations: Ethical Engineering as Inclusive Engineering

Outside of the lecture halls and late-night problem set groups, CS students could find community on campus through student organizations, the most vibrant of which, at least in my experience, was Harvard Women in Computer Science (WiCS). Established with the goal of "building a community of technical women at Harvard and beyond"[70], WiCS offers, amongst other things, mentorship programs; maintains one of the most active mailing lists for CS students; and hosts the annual WECode conference that celebrates "the accomplishments of women and historically underrepresented minorities in computer science and technology"[71].

Harvard WiCS, therefore, in my mind, embodied the idea of ethical engineering as inclusive engineering. An ethical engineer, according to the inclusivity narrative, is an engineer who actively works towards a more equitable and welcoming engineering culture in both academia and industry. In practice, this could mean that an inclusive engineer may volunteer in programs such as Girls Who Code to challenge stereotypes of who can be a software engineer[72]. Inclusive engineers may also support women and other underrepresented groups in engineering in attaining practical career skills, such as writing resumes or practicing coding interviews. Excelling in these skills is of particular importance for people who have to navigate professional spaces that were not designed for them and, in fact, often are actively hostile towards them.[11] In the workplace, ethical engineers may make a conscious effort to listen to and uplift the voices of colleagues, who may otherwise be subject to (un)conscious bias[75]. Ethical engineers may also actively

[11] The James Damore Memo[73] and stories recounted in the book Brotopia[74] being just two examples of an overall misogynistic culture in the technology industry.

work towards empathy being considered a core engineering skill that is essential to building great software in diverse, distributed, and multidisciplinary teams[76].

These efforts seem to work. Several women classmates of mine mentioned the support of the WiCS community as one of the main reasons they felt empowered to pursue and remain in a computer science major at Harvard. Interacting with faculty members who were supportive of WiCS, finding problem set partners (which can be challenging if you join the CS major after freshman year without an already established problem setting group), or finding roommates during summer internships via a crowdsourced spreadsheet were just a few examples of the community and support that WiCS fostered. From its founding in 2012 until 2015, “the percentage of female concentrators in CS has risen from roughly 1/4 to about 1/3”[77].

This success left me wondering: Will this representation translate to better outcomes of technology production more generally, and if so how? Broader discussions about the societal impact of the tech industry’s products were, at least from my vantage point, less visible among the many coffee chats with recruiters throughout the year and the corporate-sponsored sessions ("Mocktails with Bloomberg" and "Zumba with Google") on the WECode conference schedule. Harvard WiCS’s approach towards inclusion seemed, to my outsider’s gaze, to at least postpone scrutinizing the tech industry itself, mainly appealing to broaden the recruiting pipeline towards it: “The technology sector is perpetually hungry for more talent, and can put that talent to good use solving some of humanity’s most pressing and difficult problems. Getting more women involved in technology could nearly double the talent supply for this critical work”[75]. What if the pipeline should better be disrupted or rerouted?

My critique of this politics of representation warrants a number of caveats. First, it by no means applies to WiCS exclusively, but to a number of student organizations on Harvard’s campus that have a strong pre-professional focus across industries such as consulting, finance[78], and technology. The gender-neutral Harvard Computer Society, for instance, also co-hosted many corporate tech talks on campus rather than asking hard questions. Second, given how difficult it already can be for people from underrepresented minorities to stay in a computer science major and, after graduation, in the technology industry itself, it seems unfair to expect the very same people to also put their energy into pushing their organizations towards more ethical outcomes of engineering, as such efforts may come with professional risks, disproportionately so for minorities[79].

My post-graduation experience in industry bore out a similar pattern. The people who were most outspoken in discussions about the impacts of their work tended to overwhelmingly be white men in senior software roles who were U.S. citizens. Younger, non-white, non-citizen, non-engineers were much less likely to openly partake in internal discussions. What would it take to broaden the set of those who can safely participate in discussions about the downstream impacts of their work? As one friend suggested, maybe, as a simple matter of capacity, it would require those people who are currently outspoken to take on more of the emotional labor and mentorship work that women tech workers currently shoulder. Third, it should be acknowledged that the debate about the scope of ethical engineering as inclusive engineering seems to be evolving, incorporating the idea of refusal that I encountered in the History of Science department. For instance, AnitaB.org, the nonprofit that organizes the annual Grace Hopper Celebration of Women in Computing Conference, recently announced the removal of Palantir as a sponsor[80], in light of the company's work with ICE.[12] Harvard's WiCS board decided to follow suit soon thereafter, ending its sponsored partnership and vowing to "create a culture of transparency"[82] around its sponsorships.

Besides my limited personal impression of Harvard WiCS (which may not generalize to other WiCS organizations) and my experience as an instructor for Girls Who Code, I have become aware of other organizations which embrace inclusivity beyond representation in an outward-facing way. Project Include, for example, is a Toronto-based organization that focuses on offering CS education in lower income neighborhoods. Leveraging their unique perspective, participants from a subsidized housing complex, for example, developed a website to facilitate sharing of sports equipment among the building's residents, suggesting that "increas[ing] participation in the computer science field, [...] also widen[s] the scope of ideas and solutions"[83]. Another instance of inclusive engineering in a wider sense that inspired me is that of the student-led Public Interest Technology Clinic at Olin College, run by four undergraduate women engineering

[12] The decision seems to have been made in response to a Change.org petition organized by the activist organization Mijente's #NoTechForICE campaign which had garnered 364 signatures[81].

students, who integrated many of the ideas I encountered as engineering as advocacy, reflexive engineering, and refusal in an interlocking manner[84].

6 Putting It All Together

My undergraduate encounters with ethical engineering—thoughtful engineering, advocacy, reflexive engineering and refusal, and inclusive engineering—demonstrated that there are numerous paradigms, which vary in scope, partially overlap, and may in some instances even conflict with one another. Precisely because of this, I believe that putting these viewpoints into conversation could help CS students and faculty to develop a richer conception of ethical engineering.

Such a dialogue could, for instance, create space for students to discuss what they believe to be the most effective site of intervention to improve the social outcomes of technology production. Is it changing our own individual behavior as practitioners (as thoughtful engineering suggests) or altering the larger social, political, and economic structures of the tech industry (as reflexive engineering emphasizes)? Can the tech industry be reformed by working within it (as thoughtful and inclusive practitioners), or can it only meaningfully be challenged from the outside (through advocacy, reflexive engineering and refusal, and collective action)? Is the inside-outside dichotomy helpful[85]? Under which circumstances can the act of introducing technology to incrementally improve a social problem actually be detrimental to more meaningful reform[86]? Reasonable people may disagree about the answers to these questions. Some may, for example, believe that there is value in a combination of pressure from advocates on the outside who effectively set an agenda which is then taken on by inclusive, thoughtful practitioners inside the tech industry to implement. Others might be concerned about industry capture[87].

Despite these inherent tensions, what unites all of the different conceptions is their call to action for technologists to take a wider view of their responsibilities than they traditionally have. Harvard's Embedded EthiCS program and new dedicated CS ethics classes have certainly broadened the set of questions any software engineer should consider "in scope" when designing a system. Compared to the little consideration that technology ethics were otherwise given in CS classes, the productive disorientation and blurring of disciplinary lines that the program and classes have sparked are commendable achievements. In the coming years, I hope that the department can build on this initial success, and evolve its ethics curriculum to expand the scope of ethical engineering even further.

Most pressing, to me, in this respect would be for the curriculum to acknowledge that “technologies embody social relations”[88] and wrestle with its corollaries. If technical artifacts indeed have politics[89], it seems only logical that computer scientists who aim to build such artifacts in an ethical way ought to adopt a broader analysis of those social relations, of power relations in society. A CS ethics curriculum should equip students with frameworks to analyze these power relations and to reason about how the introduction of a technical system may “rearrange power,” in the words of cryptographer Philip Rogaway[90]. How, for example, does the deployment of an automated “healthcare, food stamp and cash benefit” provisioning system reconfigure the power relationships between state governments and poor people[91]? How does the introduction of a centralized ranking system for universities shift power between college applicants, university administrators, and U.S. News & World[8]? How does the introduction of a digital reputation system for ride-hailing change the distribution of power between drivers, riders, taxi companies, and regulators[92]?

Similar to how Embedded EthiCS has managed to foreground the inescapability of having to make ethical choices, as I described earlier, a curriculum that adopts a lens of power, as I suggest, would foreground the inescapability of having to make choices about how computation rearranges power. In whose favor does computer science currently tend to reconfigure power? In whose favor do students and faculty believe their field should reconfigure power, and how might they bring these two things into closer alignment? Should CS departments actively and explicitly commit their teaching activities and research agendas to the furtherance of the public good and the needs of marginalized communities[93][88]?

These questions are invariably social and political in nature. On a very practical level, CS departments may lack faculty equipped to teach courses that operate within such a scope. Luckily, scholars of political science, science and technology studies, and anthropology, amongst other fields, have been grappling with these very questions for a while. Integrating with departments such as Government and History of Science, via dedicated semester-long courses and as part of the interspersed modules, I hope, could bring to the fore the political economy of software engineering and further de-center technologists from conversations about what role technology should play in our lives. The design justice framework could serve as a model[88] for such collaborations.

Some of the insights that students would gain in a curriculum that views computation through a lens of power may be unsettling or even downright depressing. Understanding them may be contrary to the immediate (narrow, economic) self-interest of undergraduates, who could spend their precious class time in courses in which they can develop marketable skills[94]. In my personal experience, however, many of my classmates who completed Embedded EthiCS modules or enrolled in semester-long classes, found them intellectually rewarding and have since been wondering about how they might put their learnings into practice.

I personally was fortunate to join a multi-disciplinary team after graduating, and found a culture that already espoused many of the abovementioned thoughtful, reflexive, and inclusive practices. I would often hear from coworkers on other teams, however, who felt like they lacked the infrastructure for such work. Considering all the different paradigms I laid out, my own day-to-day work nonetheless undoubtedly benefited most from the efforts of external engineer advocates. Their scrutiny was what gave my coworkers and me the critical support to push internally for products to be built more responsibly. It meaningfully altered the decision calculus of executives who might otherwise have not embraced our solutions, many of which involved short-term negative effects on business outcomes. More generally, coworkers who care deeply about ethical technology production gravitated towards lines of work that were spun up in response to or anticipation of regulatory action, spaces for ethical engineering which external engineer advocates had opened for us.

I hope ethics courses and modules will more explicitly leverage the lived experience of their students, many of whom have interned in technology companies in previous summers by the time they attend such a class or module. Giving students space to talk about their experiences as tech workers and taking those as a point of departure to analyze how capitalism and technology production intersect could be worthwhile. Additionally, I believe that tech ethics courses could benefit not just from a multidisciplinary set of instructors and course material, but also from a multidisciplinary student body. Courses that span a variety of majors would also prepare students for multidisciplinary work environments that require "a mutual willingness to understand each other's languages and professional cultures"[58]. Ideally, these reflexive engineering practices will not only be useful to students, but even help CS faculty to reflect on larger questions in their department around values embedded in research agendas, hiring criteria, and industry partnerships.

I took away from my encounters with these different paradigms of ethical engineering the contours of my own, more holistic definition of an ethical engineer, who I think of as somebody who is thoughtful about the systems they build day-to-day; who does their best to anticipate and address harms and unintended consequences; who situates their work in a larger societal, economic, and political context; who uses the privileges they have been endowed with to further inclusivity; who understands that it is not their place to decide on our technological futures and instead makes their expertise available for communities and policy-makers to self-determine on the basis of human rights; and who realizes that sometimes refusal is the most constructive position to take.

References

[1] AlgorithmWatch, "AI Ethics Guidelines Global Inventory," *Algorithm Watch*, Jun. 21, 2019. https://algorithmwatch.org/en/project/ai-ethics-guidelines-global-inventory/

[2] F.-F. Li and J. Etchemendy, "Opening the Gate," *Stanford Institute for Human-Centered Artificial Intelligence,* Mar. 11, 2019. https://hai.stanford.edu/news/opening-gate (accessed Mar. 02, 2020).

[3] MIT News Office, "MIT reshapes itself to shape the future," *MIT News*, Oct. 15, 2018. http://news.mit.edu/2018/mit-reshapes-itself-stephen-schwarzman-college-of-computing-1015 (accessed Mar. 03, 2020).

[4] University of Oxford, "University announces unprecedented investment in the Humanities," Jun. 19, 2019. http://www.ox.ac.uk/news/2019-06-19-university-announces-unprecedented-investment-humanities (accessed Mar. 03, 2020).

[5] J. Metcalf, E. Moss, and D. Boyd, "Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics," *Soc. Res. Int. Q.*, vol. 86, no. 2, pp. 449–476, Aug. 2019.

[6] R. Benjamin, *Race after technology: Abolitionist tools for the new jim code*. John Wiley & Sons, 2019.

[7] S. U. Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, 2018.

[8] C. O'Neil, *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books, 2016.

[9] L. Sweeney, "Discrimination in online ad delivery," *Queue*, vol. 11, no. 3, pp. 10–29, 2013.

[10] J. Weedon, W. Nuland, and A. Stamos, "Information operations and Facebook," 2017, [Online]. Available: https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf

[11] Y. Benkler, R. Faris, and H. Roberts, *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press, 2018.

[12] J. C. Ong and J. Cabanes, "Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the Philippines," *Newton Tech4Dev Netw.*, 2018.

[13] Harvard College, *Harvard College Handbook for Students 2019-2020*. Accessed: Mar. 03, 2020. [Online]. Available: https://handbook.fas.harvard.edu/book/computer-science

[14] B. J. Grosz *et al.*, "Embedded EthiCS: integrating ethics across CS education," *Commun. ACM*, vol. 62, no. 8, pp. 54–61, 2019.

[15] C. Goodrow, "You know what's cool? A billion hours," *Official YouTube Blog*, Feb. 27, 2017. https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html (accessed Mar. 03, 2020).

[16] L. Bever, "Censorship or social responsibility? Amazon removes some books peddling vaccine misinformation.," *Washington Post*, Mar. 18, 2019.

https://www.washingtonpost.com/business/2019/03/18/censorship-or-social-responsibility-amazon-removes-some-books-peddling-vaccine-misinformation/ (accessed Mar. 03, 2020).

[17] C. Caron, "Pinterest Restricts Vaccine Search Results to Curb Spread of Misinformation," *The New York Times*, Feb. 23, 2019. Accessed: Mar. 03, 2020. [Online]. Available: https://www.nytimes.com/2019/02/23/health/pinterest-vaccination-searches.html

[18] C. Caron, "Facebook Announces Plan to Curb Vaccine Misinformation," *The New York Times*, Mar. 07, 2019. Accessed: Mar. 03, 2020. [Online]. Available: https://www.nytimes.com/2019/03/07/technology/facebook-anti-vaccine-misinformation.html

[19] C. O'Donovan and L. McDonald, "YouTube Continues To Promote Anti-Vax Videos As Facebook Prepares To Fight Medical Misinformation," *BuzzFeed News*, Feb. 20, 2019. https://www.buzzfeednews.com/article/carolineodonovan/youtube-anti-vaccination-video-recommendations (accessed Mar. 03, 2020).

[20] K. Roose, "The Making of a YouTube Radical," *The New York Times*, Jun. 08, 2019. Accessed: Mar. 03, 2020. [Online]. Available: https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html

[21] F. Doshi-Velez, *CS181: Introduction to Machine Learning*. harvard-ml-courses, 2018. Accessed: Mar. 03, 2020. [Online]. Available: https://github.com/harvard-ml-courses/cs181-lectures

[22] "Homework 3: Max-Margin, Ethics, Clustering," *GitHub*, 2019. https://github.com/harvard-ml-courses/cs181-s19-homeworks (accessed Mar. 03, 2020).

[23] K. Vredenburgh, "Machine Learning and Discrimination," Spring 2018. Accessed: Mar. 03, 2020. [Online]. Available: https://github.com/harvard-ml-courses/cs181-lectures

[24] J. Kun, "Reductions are the Mathematical Equivalent of Hacks," *Jeremy Kun on Svbtle*, May 05, 2014. https://j2kun.svbtle.com/polynomial-time-reductions-are-hacks (accessed Mar. 09, 2020).

[25] M. Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press, 2010.

[26] L. Stark and A. L. Hoffmann, “Data is the new what? Popular metaphors & professional ethics in emerging data culture,” *Journal of Cultural Analytics*, 2019.

[27] L. Winner, *Autonomous technology: Technics-out-of-control as a theme in political thought*. Mit Press, 1978.

[28] Harvard University Department of Government, “Tech Science.” https://undergrad.gov.harvard.edu/tech-science-within-government (accessed Mar. 03, 2020).

[29] R. Abebe, S. Barocas, J. Kleinberg, K. Levy, M. Raghavan, and D. G. Robinson, “Roles for computing in social change,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, pp. 252–260.

[30] M. Blaze *et al.*, “Source code review of the Sequoia voting system,” *State California’s Top Bottom Rev.*, vol. 2, 2007.

[31] C. Kurz and F. Rieger, “NEDAP-Wahlcomputer–Manipulationsmethoden an Hard- und Software,” *Inform.-Spektrum*, vol. 30, no. 5, pp. 313–321, 2007.

[32] H. Abelson *et al.*, “Keys under doormats: mandating insecurity by requiring government access to all data and communications,” *J. Cybersecurity*, vol. 1, no. 1, pp. 69–79, 2015.

[33] B. Schneier, “Should U.S. Hackers Fix Cybersecurity Holes or Exploit Them?,” *The Atlantic*, May 19, 2014. https://www.theatlantic.com/technology/archive/2014/05/should-hackers-fix-cybersecurity-holes-or-exploit-them/371197/ (accessed Mar. 03, 2020).

[34] C. Garvie, A. Bedoya, and J. Frankle, *The Perpetual Line-Up: Unregulated Police Face Recognition in America. Georgetown Law Center on Privacy & Technology*. 2016.

[35] L. Sweeney, “Simple demographics often identify people uniquely,” *Health San Franc.*, vol. 671, pp. 1–34, 2000.

[36] R. Sietmann, “CCC fordert Verbot von Wahlcomputern - Nedap wehrt Vorwürfe ab,” *heise online*, Oct. 05, 2006. https://www.heise.de/newsticker/meldung/CCC-fordert-

Verbot-von-Wahlcomputern-Nedap-wehrt-Vorwuerfe-ab-168667.html (accessed Mar. 05, 2020).

[37] J. Buolamwini and T. Gebru, “Gender shades: Intersectional accuracy disparities in commercial gender classification,” in *Conference on fairness, accountability and transparency*, 2018, pp. 77–91.

[38] U. D. of Health and H. Services, “Guidance regarding methods for de-identification of protected health information in accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule,” *US Dep. Health Hum. Serv. Wash. DC*, vol. 26, p. 2018, 2012.

[39] 2 Senat Bundesverfassungsgericht, “Bundesverfassungsgericht - Entscheidungen - Partielle Unvereinbarkeit der Bundeswahlgeräteverordnung mit Art 38 GG iVm Art 20 Abs 1, Abs 2 GG mangels Sicherstellung einer dem Grundsatz der Öffentlichkeit der Wahl entsprechenden Kontrolle - Unvereinbarkeit der Verwendung von elektronischen Wahlgeräten der Firma Nedap bei der Wahl zum 16. Deutschen Bundestag mit Art 38 GG iVm Art 20 Abs 1, Abs 2 GG,” Mar. 03, 2009. https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2009/03/cs20090303_2bvc000307.html (accessed Mar. 05, 2020).

[40] I. D. Raji and J. Buolamwini, “Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products,” in *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, Honolulu, HI, USA, Jan. 2019, pp. 429–435. doi: 10.1145/3306618.3314244.

[41] A. Khanna, “Venmo’ed: Sharing your payment data with the world,” *Technology Science*, 2015.

[42] M. Weiss, “Deepfake Bot Submissions to Federal Public Comment Websites Cannot Be Distinguished from Human Submissions,” *Technol. Sci.*, vol. 2019121801, 2019.

[43] P. Braunschweiger and K. W. Goodman, “The CITI program: an international online resource for education in human subjects protection and the responsible conduct of research,” *Acad. Med.*, vol. 82, no. 9, pp. 861–864, 2007.

[44] National Commission for the Protection of Human Subjects of Biomedicaland Behavioral Research, *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Superintendent of Documents, 1978.

[45] Nieman Lab Communications, “Harvard project to track personal data wins Knight News Challenge award,” *Harvard Gazette*, Feb. 01, 2016. https://news.harvard.edu/gazette/story/2016/02/harvard-project-to-track-personal-data-wins-knight-news-challenge-award/ (accessed Mar. 05, 2020).

[46] A. Crocker and C. Fischer, “Victory! Ruling in hiQ v. Linkedin Protects Scraping of Public Data,” *Electronic Frontier Foundation*, Sep. 10, 2019. https://www.eff.org/deeplinks/2019/09/victory-ruling-hiq-v-linkedin-protects-scraping-public-data (accessed Mar. 09, 2020).

[47] J. Taylor, “Melbourne professor quits after health department pressures her over data breach,” *The Guardian*, Mar. 07, 2020. Accessed: Mar. 09, 2020. [Online]. Available: https://www.theguardian.com/australia-news/2020/mar/08/melbourne-professor-quits-after-health-department-pressures-her-over-data-breach

[48] P. Farrell, “Research work could be criminalised under George Brandis data changes,” *The Guardian*, Sep. 29, 2016. Accessed: Mar. 09, 2020. [Online]. Available: https://www.theguardian.com/world/2016/sep/29/george-brandis-to-criminalise-re-identifying-published-government-data

[49] G. King and N. Persily, “A New Model for Industry–Academic Partnerships,” *PS Polit. Sci. Polit.*, pp. 1–7, 2018.

[50] TechCongress, “TechCongress: A Congressional Innovation Fellowship,” *TechCongress*. https://www.techcongress.io (accessed Mar. 05, 2020).

[51] B. Schneier, “Public-Interest Technology Resources,” Feb. 24, 2020. https://public-interest-tech.com/ (accessed Mar. 05, 2020).

[52] H. Marcus and B. Wilson, “Syllabus: History of Science 100. Knowing The World. An Introduction to the History of Science.”

[53] T. Jefferson, *Notes on the State of Virginia*. Penguin, 1999.

[54] N. C. M. Hartsock, “The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism,” in *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, S. Harding and M. B. Hintikka, Eds. Dordrecht: Springer Netherlands, 1983, pp. 283–310. doi: 10.1007/0-306-48017-4_15.

[55] D. Haraway, “Situated knowledges: The science question in feminism and the privilege of partial perspective,” *Fem. Stud.*, vol. 14, no. 3, pp. 575–599, 1988.

[56] M. R. Smith and L. Marx, *Does Technology Drive History?: The Dilemma of Technological Determinism*. MIT Press, 1994.

[57] J. Sadowski and E. Ongweso Jr, *This Machine Kills 32. Conceptual Hit List*. 2021. Accessed: Nov. 18, 2021. [Online]. Available: https://soundcloud.com/thismachinekillspod/32-conceptual-hit-list

[58] P. T. Robbins, “The reflexive engineer: perceptions of integrated development,” *J. Int. Dev.*, vol. 19, no. 1, pp. 99–110, 2007, doi: 10.1002/jid.1351.

[59] “Um die Ecke,” *Der Spiegel*, Oct. 17, 1982. Accessed: Nov. 18, 2021. [Online]. Available: https://www.spiegel.de/wirtschaft/um-die-ecke-a-c87f8e69-0002-0001-0000-000014351295

[60] S. Levy, “Facebook Failed the People Who Tried to Improve It,” Oct. 25, 2021. https://www.wired.com/story/facebook-papers-badge-posts-former-employees/ (accessed Nov. 18, 2021).

[61] F. Turner, *From counterculture to cyberculture: Stewart Brand, the Whole Earth Network, and the rise of digital utopianism*. University of Chicago Press, 2010.

[62] Princeton Dialogues on AI and Ethics, “Hiring By Machine,” *Princeton Dialogues on AI and Ethics*, Jan. 24, 2019. https://aiethics.princeton.edu/case-studies/case-study-pdfs/ (accessed Mar. 05, 2020).

[63] Logic Magazine, “Hack the Planet: Tega Brain on Leaks, Glitches, and Preposterous Futures,” *Logic Magazine*, Dec. 07, 2019. https://logicmag.io/nature/hack-the-planet-tega-brain-on-leaks-glitches-and-preposterous-futures/ (accessed Mar. 05, 2020).

[64] J. Penn, “AI thinks like a corporation—and that’s worrying,” *The Economist*, Nov. 26, 2018. Accessed: May 28, 2020. [Online]. Available: https://www.economist.com/open-future/2018/11/26/ai-thinks-like-a-corporation-and-thats-worrying

[65] B. Tarnoff, “From Manchester to Barcelona,” *Logic Magazine*, Dec. 07, 2019. https://logicmag.io/nature/from-manchester-to-barcelona/ (accessed Mar. 09, 2020).

[66] A. Glaser, "The Techlash Has Come to Stanford," *Slate Magazine*, Aug. 08, 2019. https://slate.com/technology/2019/08/stanford-tech-students-backlash-google-facebook-palantir.html (accessed Mar. 05, 2020).

[67] H. Natanson and D. G. Xiao, "Class of 2019 By the Numbers," 2019. https://features.thecrimson.com/2019/senior-survey/ (accessed Mar. 05, 2020).

[68] E. Goldberg, "'Techlash' Hits College Campuses," *The New York Times*, Jan. 11, 2020. Accessed: Mar. 05, 2020. [Online]. Available: https://www.nytimes.com/2020/01/11/style/college-tech-recruiting.html

[69] E. Graeff, "The Responsibility to Not Design and the Need for Citizen Professionalism," *Tech Otherwise*, May 2020, doi: 10.21428/93b2c832.c8387014.

[70] Harvard Women in Computer Science, "Mission," *Harvard Women in Computer Science*. http://www.harvardwics.com/mission (accessed Mar. 05, 2020).

[71] Harvard Women in Computer Science, "What We Do," *Harvard Women in Computer Science*. http://www.harvardwics.com/what-we-do (accessed Mar. 05, 2020).

[72] Girls Who Code, "Our Values," *girlswhocode*. https://girlswhocode.com/our-values/ (accessed Mar. 15, 2020).

[73] "Google's Ideological Echo Chamber," *Wikipedia*. Mar. 05, 2020. Accessed: Mar. 05, 2020. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Google%27s_Ideological_Echo_Chamber&oldid=943981694

[74] E. Chang, *Brotopia: Breaking Up the Boys' Club of Silicon Valley*. Portfolio, 2019.

[75] WiCS Advocacy Council, "Why We Care." http://advocacy.harvardwics.com/report/why-we-care/ (accessed Mar. 09, 2020).

[76] J. Walther, S. E. Miller, and N. W. Sochacka, "A model of empathy in engineering as a core skill, practice orientation, and professional way of being," *J. Eng. Educ.*, vol. 106, no. 1, pp. 123–148, 2017.

[77] Harvard Women in Computer Science, "History," *Harvard Women in Computer Science*. http://www.harvardwics.com/history (accessed Mar. 05, 2020).

[78] G. W. Bishai, "Protest, False Fire Alarm Disrupt Martin Shkreli's Harvard Speech," *The Harvard Crimson*, Feb. 16, 2017.

https://www.thecrimson.com/article/2017/2/16/martin-shkreli-protest/ (accessed Mar. 05, 2020).

[79] AI Now Institute, “How To Interview a Tech Company,” *Medium*, Sep. 27, 2019. https://medium.com/@AINowInstitute/how-to-interview-a-tech-company-d4cc74b436e9 (accessed Mar. 05, 2020).

[80] S. Ghaffary, “The world’s biggest women’s tech conference just dropped Palantir as a sponsor,” *Vox*, Aug. 28, 2019. https://www.vox.com/recode/2019/8/28/20837365/anita-b-grace-hopper-palantir-sponsor-worlds-biggest-womens-tech-conference-dropped (accessed Mar. 05, 2020).

[81] “AnitaB.Org: Drop Palantir as a Grace Hopper Celebration sponsor,” *Change.org*, Aug. 28, 2019. https://www.change.org/p/tell-anitab-org-to-drop-palantir-as-a-grace-hopper-celebration-sponsor (accessed Mar. 09, 2020).

[82] Harvard Women in Computer Science, “[wics-chatter] WiCS and Palantir,” Oct. 13, 2019.

[83] *Power Shift: Equity In the Tech Sector*, (Sep. 28, 2021). Accessed: Nov. 18, 2021. [Online Video]. Available: https://vimeo.com/617205979

[84] S. Chowdhary, S. Daitzman, R. Eisenbud, E. Pan, and E. Graeff, “Care and Liberation in Creating a Student-Led Public Interest Technology Clinic,” in *2020 IEEE International Symposium on Technology and Society (ISTAS)*, 2020, pp. 164–175.

[85] M. Rakowitz, S. Shukaitis, S. Rodriguez, and A. Petrossiants, “Diversity of Aesthetics Vol. 1,” *Printed Matter*, 2021. https://www.printedmatter.org/catalog/58285/ (accessed Nov. 18, 2021).

[86] B. Green, Data Science as Political Action: Grounding Data Science in a Politics of Justice, *Journal of Social Computing*, 2021.

[87] M. Whittaker, “The steep cost of capture,” *Interactions*, vol. 28, no. 6, pp. 50–55, 2021.

[88] S. Costanza-Chock, *Design justice: Community-led practices to build the worlds we need*. MIT Press, 2020.

[89] L. Winner, “Do artifacts have politics?,” *Daedalus*, pp. 121–136, 1980.

[90] P. Rogaway, “The Moral Character of Cryptographic Work.,” *IACR Cryptol. EPrint Arch.*, vol. 2015, p. 1162, 2015.

[91] V. Eubanks, *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press, 2018.

[92] V. Dubal, "A Brief History of the Gig," *Logic Magazine*, May 04, 2020. https://logicmag.io/security/a-brief-history-of-the-gig/ (accessed May 18, 2020).

[93] The Radical AI Project, "The Radical AI Project." 2020. [Online]. Available: http://radicalaiproject.org

[94] P. Rogaway, "Letter to ECS 189L students," Mar. 18, 2020. Accessed: May 10, 2020. [Online]. Available: https://web.cs.ucdavis.edu/~rogaway/classes/189/spring20/covid19.pdf