It's about power: What ethical concerns do software engineers have, and what do they (feel they can) do about them?

David Gray Widder*
dwidder@cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Derrick Zhen*
dzhen1@swarthmore.edu
Swarthmore College
Swarthmore, Pennsylvania, USA

Laura Dabbish
dabbish@cmu.edu
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

James Herbsleb
jim.herbsleb@gmail.com
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

ABSTRACT

How do software engineers identify and act on their ethical concerns? Past work examines how software practitioners navigate specific ethical principles such as "fairness", but this narrows the scope of concerns to implementing pre-specified principles. In contrast, we report self-identified ethical concerns of 115 survey respondents and 21 interviewees across five continents and in non-profit, contractor, and non-tech firms. We enumerate their concerns – military, privacy, advertising, surveillance, and the scope of their concerns – from simple bugs to questioning their industry's entire existence. We illustrate how attempts to resolve concerns are limited by factors such as personal precarity and organizational incentives. We discuss how even relatively powerful software engineers often lacked the power to resolve their ethical concerns. Our results suggest that ethics interventions must expand from helping practitioners merely *identify* issues to instead helping them build their (collective) *power* to resolve them, and that tech ethics discussions may consider broadening beyond foci on AI or Big Tech.

ACM Reference Format:
David Gray Widder, Derrick Zhen, Laura Dabbish, and James Herbsleb. 2023. It's about power: What ethical concerns do software engineers have, and what do they (feel they can) do about them?. In *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3593013.3594012

1 INTRODUCTION AND RELATED WORK

Facing public pressure and negative press [27], many large technology companies are attempting to address harms from algorithmic systems, often by instituting ethics initiatives which converge on principles such as transparency or fairness [33]. Metcalf *et al.* show how broad Silicon Valley logics cloud official ethics initiatives [44], some startup environments see ethics work as premature [64], and even some "major companies" see ethics work as "too complicated for the organization's current level of resources" [56]. A variety of interventions have been proposed to make operationalizing AI ethics easier, including fairness checklists [42], fairness toolkits [72] datasheets [20] and model cards [46]. However, some argue that the convergence around codified principles like "fairness" or "accountablity" obscures underlying political and normative disagreements [47], and there is increasing evidence for this: AI practitioners have different values than the general public in AI system design [32], and workers have different concerns than those who seek to monitor them [54]. These principles may also have discordant definitions [16, 17, 40], and others argue that principles limit scrutiny to a system's design without scrutinizing *use* [35, 70] or business decisions [24]. These concerns lead to accusations of "ethics washing" [26, 66]: where companies put forward voluntary principles to burnish their reputation and avoid regulation [52], without changing their behavior [73].

Given that software practitioners have some agency in how to develop these systems [53], research examines their needs and behavior as they seek to build ethical systems. For example, past work questions the effect of ethical codes [22] on software engineers' ethical decision making [43]. In machine learning specifically, Holstein *et al.* examine practitioners' challenges in developing fair systems [30], Madaio *et al.* examine practitioners' challenges in using disaggregated evaluations to assess system fairness [41], and Veale *et al.* examine the needs of public sector practitioners in ensuring fairness and accountability in high stakes systems [65]. However, as discussed above, studies which focus on "fairness" or "accountablity" may impose a narrow scope of scrutiny and thus foreclose on wider concerns, and many software practitioners work at smaller companies that may not have official ethics initiatives. Given these concerns, we surveyed 115 and interviewed 21 software engineers about their *self-identified* ethical concerns, as opposed to concerns identified using codified ethical principles, toolkits, or codes, to answer: **RQ1: What are software engineers' ethical concerns?** With this open scope, we discuss both the kinds of concerns our participants raise – military, privacy, advertising, surveillance, and others – but also examine the scope of their concerns: ranging from concerns about bugs which can me more easily fixed, to wider concerns questioning their company's *raison d'être*.

Others study what happens after tech workers of various stripes develop ethical concerns. Whereas some AI practitioners engage

*The first two authors contributed equally to this research.

FAccT '23, June 12–15, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0192-4/23/06.
https://doi.org/10.1145/3593013.3594012

in high-profile activism [5], Madaio *et al.* find that others advocate less strongly for fairness issues due to career concerns [42], and Richmond Wong shows how User Experience practitioners employ softer tactics of resistance [71]. Some study how relatively less powerful gig workers resist opaque algorithmic evaluation [55] and use online forums to seek to understand algorithmic management they are subjected to [38], and how crowd workers engage in collective action [60]. Nedzhvetskaya and Tan collected examples of blue and white collar tech worker engaging in collective action [49], and discuss how workers claim they ought to have a role in AI ethics governance [50]. Similarly, after collecting software engineer's ethical concerns, we investigate how they respond: **RQ2: What happens when software engineers develop ethical concerns?** We report on a broad variety of actions participants take – from proposing technical fixes, to negotiating within organizational incentives, to resigning in protest – and on the psychological toll that these actions lead to.

Power is increasingly recognized as a central factor when differently situated actors raise tech ethics concerns. For example, recent work examines power asymmetries as students resist algorithmically lowered grades [6], and how software engineers see themselves as less powerful "mediators between powerful bodies" [53]. Others position software engineers as powerful actors in AI ethics given high demand for their labor [14]. A recent critical analysis of AI fairness toolkits find that they frequently ignore organizational power dynamics [72], and a recent study reviewing work from FAccT and a related venue finds that future work ought to attend to "structural and historical power asymmetries" [8]. In line with this call, we examine contingencies of software engineers' power as they raise ethics concerns, with an eye towards how these contingencies explicitly factor into the actions they choose to take: **RQ3: What affects software engineers' power to resolve their concerns?** We find that financial and immigration precarity, workplace culture, and organizational incentives constrain participants' power to see their concerns resolved.

After detailing our survey and interviewing methods, successive sections answer each research question. We then discuss the implications of our work: that future tech ethics research ought to turn from helping spot issues to helping practitioners build their power to actually fix them, and we finally question the foci on AI or Big Tech in tech ethics discourse.

2 METHODS AND PARTICIPANTS

We seek to understand practitioners' self-identified ethical concerns and how they navigate them. Therefore, we imposed no *a-priori* definition of ethics, nor do we seek to reach a singular definition in our work: instead, in our survey instrument, we use an open-ended framing to ask survey respondents if they have "ever had ethical concerns with a software system they were asked to contribute to" and actions they took, resolutions, factors which made raising concerns harder or easier, and an invitation to an optional follow up interview (instrument in Appendix 8.1). To recruit a broad sample, we recruited using diverse methods including posts to Twitter, software engineering message boards, software-ethics focused messaging channels, the popular StackOverflow programming Q&A site's blog; and in person at a developer meetup.

The survey was open for 87 days from May to August 2022, and received 115 responses. 90 survey respondents were employed full-time, 15 were employed part-time or as contractors and 10 were not currently employed. 13 respondents worked at very small firms (<10 employees), 29 respondents at small firms (10-99), 31 at medium sized firms (100-999) and 35 at large firms (1000+), 7 did not report the size of their firms. Respondents were relatively experienced, reporting a mean of 17 years of experience coding (med. = 15, min. = 4, max. = 46 years). Respondents spanned six continents: 68 participants lived in North America, 34 in Europe, 4 in Australia, 4 in Asia, 3 in South America, and 2 in Africa. 80 participants identified as male, 10 as female, 6 as nonbinary or nonconforming, 5 self-described and 14 preferred not to answer. 21 survey respondents participated in the optional follow-up interview (demographics in Appendix 8.3). We conducted semi-structured [68] teleconference interviews to collect a detailed order of events as practitioners navigated their concerns, to probe into their recollections of their thoughts and feelings, about factors affecting their agency and power to see their concerns resolved, and their work since (see sample protocol in Appendix 8.2). Interviews were recorded with participant consent and IRB approval and lasted between 21 and 73 minutes (mean, med: 41 min.).

We analyzed survey and interview responses sequentially. The first two authors performed an open qualitative card sort [75] on survey responses, negotiating disagreements and adjusting categories as necessary. On interview transcripts, the first two authors performed two rounds of iterative [67] thematic analysis on this data [10]: an initial round of open coding, and then the development and application of a closed coding frame. Our study *makes use of* self-selection [61] to recruit those with self-identified ethical concerns without any pre-ordained scope, but therefore our results *do not* support general claims, such as the overall prevalence of a given concern. Interviews and surveys collect self-reported experiences, risking social desirability [48], and hindsight biases [28]. Instruments were in English, a widely-spoken language for intercultural engineering communication [58], but our findings may not generalize to software engineers working in other languages.

3 RQ1: WHAT ARE SOFTWARE ENGINEERS' ETHICAL CONCERNS?

We answer this research question in two ways: firstly, explaining the *kinds* of ethical concerns raised in our survey most frequently as surfaced by our card sort. Secondly, as a spectrum illustrating the different *scopes* of practitioners' concerns, according to how much of their organization's priorities their concern calls into question.

3.1 Kinds of Ethical Concerns

Military: 17 practitioners wrote about concerns related to military applications of their work. Of practitioners who disclosed details about the systems they worked on, the most common concerning system was autonomous drone navigation software (n=5): *"Work on autonomous drone visual navigation in a GPS-denied environment" (S98).* Other respondents develop training software: *"software in support of simulations used to train US warfighters" (S161)*, logistics software for military organizations: *"I contributed to the development of a proprietary platform-as-a-service used in defense contracts" (S174)*

and engineering support software: *"an analysis tool that automatically finds errors in aeroplane jet engines." (S188)* Respondents were primarily concerned that their work would physically injure or kill others: *"I was concerned whether the software I was contributing to was being used to harm innocent civilians or infringe on human rights" (S174)*, but several raised broader ideological concerns with the militaries who used their systems, one asking: *"am I indirectly contributing to the ills of imperialism?" (S161)*

Privacy: 14 practitioners expressed concerns relating to privacy, most commonly about geotracking (n=4), one saying they *"grab[ed] geolocation data from customers [but] our product doesn't use geolocations." (S67)*. Others were concerned about *"stor[ing] user keystrokes in a signup form to a marketing and analytics platform before the user actually submitted the form" (S272)*, scraping social media profiles *"as part of additional information to include when making loan decision." (S114)*, and privacy involved in data labeling on private footage: *"contractors [were] to label hundreds of thousands of [home security] video clips" (S26)*, or requiring personal data *"not necessary to have for the task at hand" (S69)*.

Advertising: 13 survey respondents reported concerns about advertising. Practitioners were concerned about building spam email systems, or *"bypass [spam] prevention measures" (S139)*, concerned that spam breached customers' privacy (S86), delegitimized email marketing (S139) and did not do good in the world (S230). One respondent wrote that being asked to *"develop a computer vision system that accurately classifies someone's demographics for customer segmentation marketing" (S78)* as something he believed to be inherently racist and sexist. Other practitioners wrote about implementing dishonest interfaces to *"push users to buy something because stock was "almost out"" (S2)* when in fact it was not, helping to air ads that were *" degrading toward women" (S22)*, and about advertising *"scummy for-profit schools." (S102)*

Surveillance: 11 respondents described being asked to contribute to systems used to surveil workers or citizens. Four respondents recounted concerns about working on existing workplace surveillance and algorithmic management (*i.e.*, [38]) systems, such as *"observing how well grocery stockers stayed on task" (S82)*. Their concerns included *"overwork [and] anxiety" (S82)*, that it might be *"illegal to measure employees' pee time" (P74)*, and that *"low sales numbers" (m)*ight be used to unjustly *"fire employee[s]." (S10)* Other practitioners were invited to work on surveillance systems for governments. One interviewee (I14) was asked to architect an intelligence gathering platform for a foreign government. Another respondent made improvements to an existing telecom surveillance system (S13). Other practitioners did not build surveillance systems directly, but were worried their system might be used as such downstream: *"the big problem was that I didn't see a way or a use case, where [facial] identification would be used in a non-ethically problematic way. So those would be at frontiers, at airports, identification in police stations." (I14)*

Environment, Labor Displacement, Inequality, and others: Categories of concerns expressed by less than 10 practitioners included environmental impact (n=4) *"monitoring system for agropecuary [livestock] business [which] is highly damaging to the environment" (S21)*, labor displacement (n=3) *"I thought the software system could very well put some people out of a job" (S201)*, and exacerbating inequality (n=3) *"statistically, there's no way they could do this without some form of systemic discrimination." (S44)* Other harms cited included overcharging customers (S54), contributing to addicting products (S150, S174), cryptocurrency as multi-level marketing (S70), inaccessibility of software (S50), jeopardizing healthcare outcomes (S66), legality (S115, S95), botnets (S133), implementing dark patterns (*i.e.*, [23]) (S100), autonomous vehicle safety (S118), and political manipulation (S104). Some had concerns with the software development process itself: using vulnerable frameworks (S143, S39), underpaid data labelers (S26), or closed-source software (S106).

3.2 Scope of concern: concerned with a bug, or your whole industry?

We also found that ethical concerns varied wildly in *scope*: varying in how much the organization's goals or priorities a given ethical concern questions. While they overlap, we illustrate this using four scopes of concern: those arising from bugs, intentional features, whole products, and finally concerns which question their organization's *raison d'être*. Scope affected outcomes: concerns questioning entrenched organizational goals were harder to resolve (see Sec. 5.3), and affected the kinds actions practitioners took (Sec. 4).

Bugs: Some practitioners described fixing bugs as their core ethical obligation, one saying: *"for a software developer, [software] quality is the core of ethics. Because if your product is unreliable, then your representations about the product are probably unethical." (I17)*. In some cases, proposing to fix bugs is uncontroversial, since maintaining intended functionality is often within an organization's best interest. For example, when a practitioner raised concerns about a bug in construction crane safety, they the practitioner described how this was enthusiastically received and resolved: *"there were a lot of really high profile accidents with lifting cranes [...] Everybody was really super on edge about making sure that our simulations were correct. [...] And so when I brought that issue up [...] they did a big investigation and found out that it was a data entry error." (I10)* However, organizational incentives can instead stifle practitioners' efforts to identify, fix and prevent bugs. For instance, one respondent felt non-technical firms tend not to invest in code maintenance as long as the software is minimally functional: *"non-tech companies [...] just care about business continuity" (S81)* Another interviewee explained how cost cutting at his consulting firm made it difficult to do work of acceptable quality.

A specific feature: Unlike bugs, features were intentional: practitioners were directed to implement them by their manager or client, and therefore questioning them often required more directly questioning their organizations' objectives. For example, one interviewee was asked to implement a feature that would round down GPS coordinates on properties being evaluated for insurance eligibility, which *"would have denied people access to certain types of insurance." (I10)* Another interviewee working on workplace compliance software reported that his boss asked him to implement a feature that he felt was privacy invasive: *"My boss [said] we need to put in a thing on the app so that we can see where people are all the time. And I told him [...] most of the people install it on their personal phone." (I6)* Other concerns arose when practitioners were disallowed from implementing features they felt were ethically important. For example, an interviewee developed ethical concerns

about how her product may be exclusionary: *"A really famous VR software at the time, had done inclusivity in terms of the color of the skin [...] and allowing for people with one hand to operate it. [...] I brought it up as an option" (I11)*, but this was not pursued and she was told *"well, nobody asked for it." (I11)*

An entire product: Practitioners also surfaced ethical concerns about entire products, or, as consultants, entire contracts they were assigned to. When respondents had concerns about a product's very existence, many felt concern could only be resolved if the product is shut down or dramatically altered. One contractor at a marketing consulting firm was assigned to develop a customer segmentation model, to help their client profit from high interest loans by: *"find[ing] customers that were likely to [...] take on unsustainable amounts of debt." (I5)* In this case, changes to the implementation of the product would not reconcile the practitioner's concern that building a product to sell "unsustainable" loans was unethical. Another interviewee reported being assigned onto a project to make improvements on telecom software which he suspected was being used for telecom surveillance: *"One of the main managers mentioned that the their main client for the device at the time was AT&T. [...] based on what the device was doing, they figured [...] the main use case [was] NSA tracking." (I13)* In this case, the practitioner's concern was with misuse of the product he was working on, which could not be resolved until the product was terminated, or its core use cases rethought.

An organization's *raison d'être:* Finally, some practitioners reported concerns with their organization's or industry's goals or business practices. Many practitioners were concerned that their work was used for military purposes, constituting the most common concern type. These included concerns of direct harm, such as *"the software I was contributing to was being used to harm innocent civilians" (S174)*, but also ideological issues, one pondering *"am I indirectly contributing to the ills of imperialism?" (S161)* One practitioner cited his newly-held Buddhist faith as the origin of his concerns that working in the "weapons domain" at all is *"really not good karmically" (I1)*, later reflecting that *"if you pay attention to what was going on, like in the wars, it doesn't have to be so esoteric as like Buddhist precepts." (I1)*

One interviewee, working at a fintech firm, felt his work *"preventing [fraudulent use] was not really an ethical challenge. The issue was more than the company as a whole, the business model [...] It was, you know, payday lending." (I2)* In this case, the interviewee felt was concerned about the very reason the company existed, reflecting that this made raising any concerns feel futile: *"you're actually asking to shut down the business. [...] you might as well say to the founders, like, 'Hey, either you shut down or I'm leaving', and they'll be like, 'Alright, leave, I guess.' It's not really a concern you can raise." (I2)* Even firms that offer services instead of products can be held to this level of scrutiny; as one practitioner held that their consulting firm's willingness to do business with shady clients comprised a core part of their business model: *"The company [...] does a fair amount of work for [...] oil companies, [...] firearms, [...] British American Tobacco [...] not exactly paragons of morality." (I5).*

4 RQ2: WHAT HAPPENS WHEN SOFTWARE ENGINEERS DEVELOP ETHICAL CONCERNS?

4.1 Technical Solutions

Some practitioners proposed technical solutions — changes in the functionality or design of a system through code modifications — in an attempt to mitigate potential harms. Technical solutions work best on *Bug* and *Feature*-scoped concerns, because harms resulting from the core purpose of a product or the business practices of an organization (*i.e.*, those later in Sec. 3.2) are not able to be resolved through changes to system implementation.

Furthermore, even when practitioners see opportunities for technical solutions, their actual implementation depends on management agreeing that perceived harms are important enough to warrant dedicating resources to fix them. For example, both interviewees I6 and I10 (whose concerns were summarized in 3.2) came up with technical solutions that would have resolved their concerns, which were dismissed by management. Interviewee I6 came up with a design affordance to minimize the privacy concern he had about employee location tracking: *"if you really desperately wanted to [...] see where each person is on site, [...] we could geofence the site [...] If they're not in [the site], the tracking is off." (I6).* While management was sympathetic to his privacy concerns – *"[my manager] agreed [...] we cannot monitor people's comings and goings" (I6)*, his geofencing solution was ultimately rejected due to resource constraints: *"he blatantly told me that's too much work. And he's not signing off on that." (I6).* Interviewee I10 proposed a solution to avoid erroneously denying people insurance coverage due to GPS rounding errors, suggesting *"we [could] have a three value response [the third being] 'maybe need to check further if it was right on the boundary"' (I10).* However, what the practitioner had experienced as a serious concern *"people need flood insurance for their houses [...] I had been victim to flooding and lost a bunch of my stuff" (I10)* was a non-issue for the client: *"the client cut me off and told me she didn't care and that [...] I just needed to do it." (I10)* He was later *"dressed [...] down for speaking out of turn with the client" (I10)*, and the manager *"threatened to fire me if I didn't do the work." (I10)*

4.2 Negotiating within organizational incentives

Practitioners also sought to resolve ethical concerns by convincing decision-makers like engineering or product managers that harms are serious enough to warrant action. Often times, this involves phrasing ethical concerns in terms of their effects on organizational incentives such as profit or product success.

For example, one ML researcher concerned about his project's use of facial identification (*i.e.*, who is in this picture?) reported successfully pivoting the direction of his project to facial verification (*i.e.*, are these two pictures the same person?). He raised ethical concerns about downstream uses like bias and surveillance to his management, but couched these within organizational incentives to pursue an easier and more achievable project (verification) instead of a more difficult one (identification): *"because we were understaffed [I said] ' [...] we don't have the resources to do it."' (I14)* Another interviewee, who had ethical concerns about improper employer vetting in a job matching application he helped develop, described

attempting to get senior management to shutter the project by appealing to the organization's core values. *"[I said] we either need to invest more money into understanding what is going on here [...] or we need to pump the brakes [...] I was quoting, you know, our organization's code of ethics and stuff like that." (I9)*

The likelihood of ethics negotiations succeeding are, as one practitioner puts it, *"entirely [dependent] on the organization and your ability to talk to people and [...] capture hearts and minds." (I2)* A practitioner's ability to affect change internally through "rocking the boat" relates to the broader work Debra Meyerson has done on "tempered radicals" [45] — leaders who leverage their status within organizations to promote their own values and ideals. The approach of affecting change from atop the corporate ladder was also suggested by one of our interviewees: *"[you could] work your way into a leadership position, and then start making different kinds of ideas" (I5)*. However, they acknowledged the fraught existence of individuals attempting both conformity and rebellion: *"you'd have to both hold on to your ideals [...] And at the same time, be willing to compromise your ideals quite heavily in order to work your way into a leadership position in the first place." (I5)*

4.3 Refusal

One common action respondents reported was refusing to work on the task they found unethical. Refusals took on various forms, the first being "quiet quitting" – reducing one's productivity. One practitioner who was asked to build a system to bypass spam filters wrote: *"I purposefully created a poor implementation and did not dedicate very much energy to make a working solution." (S49)*. Another respondent wrote that they *"pretended to complete the task but didn't" (S62)*. We found that the tactic of "quiet quitting" emerged from a feeling of powerlessness to affect change within organizations, and as a result is often accompanied by searching for other jobs (see 4.4). One practitioner who reported reducing productivity felt that it was impossible to resolve their concerns internally, since the product they were concerned about was already in production: *"I don't think I had any power in this dynamic because [the product] was already deployed. This was just like a minor upgrades [to] make it more usable." (I13)* Since the practitioner saw little utility in pursuing a resolution internally, they *"reduced productivity to a minimum and found another job" (S6)*.

A handful of respondents reported seeking reassignment to a different project. These practitioners removed themselves personally from the concerning project, but did not attempt to use their leverage to shut the project down: *"I was given another project to work on. I didn't kill the project, but I also didn't contribute to it." (S19)*. Reassignment is typically only possible at organizations with many product lines or clients, and practitioners felt they needed seniority to ask for reassignment, as one described: *"My seniority and wide swathe of other projects to choose from" (S19)* made securing a reassignment easier. One participant described a policy to make it easy to seek reassignment on ethical grounds: *"We had a policy at the company that nobody has to take part in any software projects involving military use." (S95)*

Other practitioners delivered ultimatums to management – putting their job on the line and making it clear that they would quit unless their concern was addressed, with mixed results dependent on their leverage and the scope of their concerns. One practitioner working on hospital software was concerned that the rushed rollout of an update would jeopardize patient outcomes. In response, he raised the concern to management forcefully: *"I looked that manager in the eye and I said: you are going to have to write me up or fire me, but I'm not doing it. I'm not going to put patients' lives at risk, because you've got a pile of money sitting on the table." (I12)*. This confrontation resulted in management stepping back and reassessing the necessity of the update. One participant suggested that ultimatums can be a wake up call for management, forcing them to take seriously harms they may have ignored in the past: *"maybe [leadership] didn't know how the individuals in the org felt. And then, individuals in the org might raise a stink. And sometimes that leads to some work being paused or just like not being done." (I7)* However, they suggest that the effectiveness of an ultimatum is highly dependent on how much leverage a practitioner has, and that collective ultimatums tend to be more effective: *"if it looks like we're gonna lose a big chunk of employees, [management] might say, we can't afford that [...] it kind of depends on the individual, whether you have leverage over leadership." (I7)*.

Resignation is typically a last resort: practitioners resign after their technical solutions or compromises are rejected (I6, I10); when escalations go sour: *"he threatened to fire me if I didn't do the work. And that's when I decided I would just quit." (I10)*; or when they lose faith that the ethical concern can be resolved internally: *"Raised concerns with executives. Started ethics discussion group among employees. Left the company after seeing no progress." (S53)*. Resignation allowed participants to put distance between themselves and the projects they deemed harmful, but they often reported this as bittersweet: in resigning, they relinquish control over development of the harmful system, as another developer is often hired on and progress resumes. One survey respondent lamented this, saying his concerns were not resolved because *"the company hired someone else. [...] I felt that I would have been in a better position ethically if I had taken the contract and had done a bad job of it." (S79)* However, in some cases, the resignation of a crucial developer in an already precarious project can terminate the project. One participant reasoned that their departure likely doomed the project: *"I was also the only one who had any serious level of software development competence [...] they generally struggled with deploying the existing models [...] so I can't imagine that they would have deployed it." (I5)* In another instance, a contract worker heard that his client canceled the project he worked on after his resignation, reflecting: *"[... quitting] can give the client cold feet on the project, it makes it look like the consulting firm is incapable of managing the project. So [...] they're likely to just cancel the project completely." (I6)*

4.4 Feet voting: "This work doesn't get done without us"

The strategy of "feet voting" describes the proactive actions practitioners took to align their employment decisions with their ethical views (such as career planning), in contrast to reactively refusing assignments or quitting jobs due to an unresolved concern. The most common action reported in this category was refusing offers of employment. Either turning down a job offer: *"I rejected the offer" (S47)*, dropping out of the interview pipeline: *"I decided to*

not continue interviewing with said job" (S45), or deciding not to apply to a position: *"Ignore the job advert" (S82)*. Many saw turning down employment to be easier than resigning, but others lamented passing up lucrative jobs: *"[Anything that made it feel harder to act?] Just the big bag of money." (I17)* or interesting projects: *"I love game development, but I don't like to work for a company that does business in gambling." (I150)*

Some practitioners with concerns about their previous industrys' *raison d'être* went to great lengths to transition to another industry. But past experience makes this difficult, as one participant trying to transition away from developing war-fighting simulations said: *"It's difficult because my experience in this industry makes me most attractive to other companies working in the same industry." (S88)* A different practitioner found it necessary to move to an entirely different state to find opportunities he was ethically aligned with: *"I realized, well, if I'm going to stay in this area, like the odds of me at some point, working [...] on defense contracts are pretty high. [...] I'm being kind of a picky applicant on what companies I'll work for. And if I really want to do that, then I might have to consider moving [...]" (I1)* He also described being more intentional in screening potential employers for red flags: *"I realized you really have to look at like the ethics of the corporation, like, as part of your interviewing process [for example, in the interview] I just asked about the details of the project [...] what space they were in, what type of product they were selling, that sort of thing." (I1)*

One practitioner argued that the favorable software engineering job market implies a unique ethical responsibility: *"Even like the 2008 financial crash [...] every software developer I knew still had work. Even if the job they had disappeared, they had a new one within a week or two. [...] I think software development is incredibly resilient against recession [...] that's why we have a responsibility to be sticks in the mud about ethics. This work doesn't get done without without us." (I10)* However, not all participants felt this way, and we discuss feelings of precarity in section 5.1.

Collective bargaining and tech worker boycotts are instances of feet voting at scale, in which practitioners collectively withhold labor from organizations they had ethical concerns with. These tactics have grown in prominence at large tech firms [37]. However, among the practitioners in our study, only one interviewee raised this *"the company would have to be pushed and that'd have to be either externally through [...] legislation or similar tools, or just public opprobrium or internally through unionization" (I5)*, mentioning that *"I did attempt to do a bit of [union] organizing work. But unfortunately, I was doing that alone." (I5)*

4.5 Leveraging legal systems

One practitioner we spoke to attempted to collect information to raise his concern with law enforcement but did not ultimately go through with it: *"I knew [...] they were going to have to start skirting rules right from the start. So, so yeah, I asked for all of the requirements, documents, anything you could give me to help me understand how to build such a system [...] My intention was just to walk into the FBI." (I17)* Another interviewee echoed this idea, saying that for harms that call into question the *raison d'être* of the entire organization (see Section 5.3), external enforcement was the sole option: *"if you do have a concern, you should take it up with the legislators or the courts." (I2)*

Practitioners who maintain open source software can also leverage laws around software licensing to prevent misuse. For instance, one practitioner personally opted to use a "copyleft" (*i.e.*, [70]) license in order to limit downstream harms of OS agricultural software they created, but conceded that it was unlikely that they would have the resources for costly litigation to enforce them. In discussing the efficacy of their action, they compared the process of choosing a license to what they saw as the small and easy yet important effect of voting as a way to effect change: *"it's a small, little one time thing you can do, that probably won't help you. But, but if it does help you, it is huge. And it only took two minutes of your time to set in place, and it's there for years, you know, that you may need to fall back on that if that's your only line of defense." (I19)*

4.6 The psychological toll of raising concerns

Practitioners reported experiencing anxiety, depression and isolation throughout the process of identifying and raising ethical concerns. The process of raising ethical concerns to an employer was stressful, especially for full-time employees, for whom their organizations are their sole benefactors. One practitioner writes: *"it terrified me to confront an 'authority' figure, especially one who was the source of my financial well-being." (S62)* Another practitioner described raising ethical concerns with a client as: *"one of the most terrifying moments in my life." (I10)* The aftermath of a failed escalation can also seriously affect practitioners' mental health, as one interviewee recalled: *"I spent a good few weeks lying in my bed [with] serious depression [...] I didn't want to leave my apartment [... I] just couldn't face [...] checking work emails." (I5)* After his concerns were dismissed by both the client and his direct manager, another described *"It gave me a lot of anxiety and depression. [...] And it kind of made me cynical [... I] approached most new working situations [...] trying to not get too involved [...]just so that it would be easier to cut and run, if somebody asked me to do something unethical." (I10)*

Practitioners also reported that just having an ethical concern at all was distressing. One interviewee quit multiple jobs over ethical concerns, recounting *"I was so distraught over what I was being asked to do, I threw up in the parking lot before going into work." (I10)* Another interviewee spoke about the alienating effect of being the only person in the office with an ethical concern: *"[I felt] kind of like an outcast" (I1)*, and another survey respondent suggested that raising concerns could lead to hostility: *"I do not want to judge, or be judged, by colleagues for my views. Without care, such discussions can lead to a hostile work environment." (S38)* However, others circulated concerns among peers in order to feel less isolated in their concerns. One interviewee leveraged their organization's employee directory and intranet to *"find other people who cared about the same things" (I4)* and start ethics reading and discussion groups. Looking back, they reflected: *"Finding community in the ethical AI space made me feel so much more grounded." (S14)*

5 RQ3: WHAT AFFECTS SOFTWARE ENGINEERS' ABILITY TO RESOLVE THEIR CONCERNS?

In this section, we discuss personal and organizational factors which affect practitioners' ability to see their concerns satisfactorily resolved, including financial and immigration precarity, company culture, and organizational incentives.

5.1 Financial and Immigration Precarity

While some software engineers felt comfortable turning down jobs (Sec. 4.4) or quitting their current jobs (Sec. 4.3) over ethical concerns, many practitioners expressed financial limitations on their power to act on their concerns. One explained how concerns over precarity took priority over ethics: *"Any kind of precarity will make your weigh your ethics less, right? [...] having a family, having dependents who can't support themselves, [...] medical conditions [and given this] you kind of are able to talk yourself into, hey, [...] I don't really have a choice." (I2)* When asked about anything that made it harder to act, survey respondents echoed this: *"The need to provide a living for me and my family, together with high prices" (S82)*, *"Reliance on the job to survive" (S8)*, and simply: *"Money." (S77)* Survey respondents also cited financial stability making it easier to act: *"I was single, didn't have a lot of debt" (S20)*,*"I had a decent savings and could afford to drop the client." (S62)* One interviewee described a stark example: *"aside from [the ethical concern...] my father had passed, and so I got some life insurance money [...] so I didn't necessarily need the paycheck anymore." (I21)* Support networks mitigate precarity: *"[My parents] said [...] they would help [me not] get put out on the street." (I10)*, but so does lacking dependents to support: *"I'm only supporting myself." (I11)*

Precarity from employment-based immigration visas (*e.g.*, US H1B visas [21]) also influenced whether practitioners decide to take action, one interviewee making clear he would only ever leave a job if he had another opportunity lined up. He said a semi-permanent state of precarity leaves immigrants less freedom act on their ethical scruples: *"Indians on H1Bs [often] need to find something [a job] within a very short period of time or actually have to leave the country. And when that happens, you end up taking whatever is available." (I2)* Practitioners were also worried about blacklisting, as one stated fear over *"getting [...] bad recommendations from former employers." (I10)* One interviewee described being blacklisted after raising an ethics concern: *"[the director] sort of ended it with like [...] I can't fire you. Because you're in contract. But like, know this: the aid sector is small. And your career here is like pretty much over." (I9)*

5.2 Workplace Culture

Respondents described how their organization's culture – including norms, expected practices, and communication styles – affected their willingness to raise concerns. For example, participants cited *"trust and respect [and] a common goal" (S71)* and an *"'Open door' policy [...] easy to get 1-on-1 time with execs" (S52)* as things that made it easier to act on concerns. However, more respondents described *"hostile" (S72)*, *"authoritarian/passive aggressive management style [...] hierarchical culture" (S110)*, *"suggestions from higher-ups that ethics discussions were a waste of time." (S52)* as things making it more difficult to act on their concerns. Interviewees expanded on this, one (I10) contrasted his two consulting experiences: the first where he worked in cubicles *"in a building full of thousands of people and feel lonely" (I10)*, and the second where he *"had good rapport"* and *"trusted"* his client and therefore felt *"safe"* to bring up his concerns.

However, some participants said "friendly" cultures made it *harder* to raise concerns. For example, survey respondents recalled that because *"The boss was a friendly chap" (S104)* or *"bonding attempts from the owners" (S68)* made it harder to raise concerns. One interviewee said that remote work meant fewer social ties, which made it easier to escalate his concerns: *"This remote way of working [...] helped me to create [...] this disconnection with the manager [which] helped me to say [...] I care less about your opinion on this." (I14)*

5.3 Organizational Incentives

Participants demonstrated an acute awareness of organizational incentives, and used them to reason about their power to act on their ethical concerns. Profit motives lead to ethical concerns, as many survey respondents identified explicitly: *"features were implemented to earn money by any means necessary" (S69)* or *"they were selling geolocation data because it's worth a lot of money." (S67)* One interviewee said that financial struggles lead to "uncomfortable" tradeoffs: *"between the choice of closing the business [versus] doing something uncomfortable, almost everyone chooses to do something uncomfortable," (I7)* recalling that a previous employer sold user data to advertisers when *"scrambling [to find] some new revenue stream?" (I7)* As seen in Section 4.2, practitioners couched ethical concerns within organizational incentives to gain support. Consequently, one interviewee described how "ethics wins" were not about an *"ethical concern, but a marketing concern, to be honest. And the way that incentives align." (I2)*

Other practitioners suggested that it is easier to resolve ethics concerns at government agencies and nonprofits, as one interviewee who had recently transitioned into public service described: *"you're pursuing your goods beyond profit, right? [...] versus 'we want to make money'." (I5)* One interviewee doing software engineering at a public university described how state funding shaped project priorities, at least in the ideal: *"If the companies paid us, I guess the situation would have been a bit different. But we [wanted] to work in the best interest of the people [...] we are paid by the people's tax money." (I20)*, but another academic described how pressures to publish led to his concerns about research integrity.

Multiple contractors and consultants described needing to compromise ethics to appease clients. One interviewee who resigned from a consulting project after being asked to do something illegal said: *"Your interactions with the client weigh very heavily on future decisions for future engagements and contracts. So there's a lot of pressure [...] to get along with the client. [...] If your client asks you to do something you don't want to do, too bad." (I12)* This appeared especially pronounced at financially precarious firms who felt the need to *"act on clients' whim[s]" (I71)* , or non-profits who feel accountable to donors rather than beneficiaries, where donors may instead have less beneficent geopolitical interests: *"the goals of the [project], are largely to keep refugees [...] in the Middle East. So they don't affect people in Europe." (I9)*

6 DISCUSSION: IT'S NOT ABOUT SPOTTING ISSUES, IT'S ABOUT HAVING *POWER* TO RESOLVE THEM

Identifying these concerns is only half of the struggle, and an unfulfilling one without the ability to ensure they are resolved. Despite recent layoffs [34], software engineers are relatively highly paid, mobile, in-demand and therefore relatively *powerful* [14] – yet our work shows that power is still a limiting factor in our participants' ability to ensure their concerns are resolved. In this section, we discuss the centrality of power in raising and resolving ethical concerns, and implications for future tech ethics research, interventions, education, and activism.

6.1 The power to declare an "ethics bug" and dedicate resources to fix it

Smaller ethical concerns, which were often described as "bugs", represent scopes of concern where a technical fix is possible, at least in theory. Ethics interventions such as toolkits [39, 57], checklists [42], principles [33], and education [18] are often designed to help practitioners identify issues, and flag them to others using artifacts such as model cards [46] or datasheets [20]. However, these interventions practically depend on practitioners having the power to dedicate resources, make design changes, or otherwise fix concerns these interventions may help identify. Without this power, these interventions risk being insufficient at best. At worst, such interventions risk limiting critique to the narrow scope of system design thereby allowing companies to avoid scrutiny of business practices [24], enforceable regulations [52, 66], or fitting into a simple narrative where morally unimaginative engineers are the core problem and training to find ethical issues the solution [76]. Our work shows what happens after practitioners identify concerns without these ethics interventions – and discover severe limits on their power to affect change as they attempt to resolve their concerns.

Similarly to how our interviewee cited academic papers on inclusivity in VR to legitimize her concerns when raising them to her team in Section 3.2, other work suggests that fairness checklists may "empower [. . .] individual advocates" [42], and other tools may enable "uncomfortable design discussions" [29] about gender bias in software design [12]. These tools legitimize ethics concerns, in part by framing them more palatably as improvements to a product (*i.e.*, as fixing "bugs") to improve its chances of success [31, 71], as some of our participants couched their ethics concerns within organizational incentives (see Sec. 4.2) and are occasionally successful (*e.g.*, crane software in Sec. 3.2).

However, we show that even when less-threatening, narrowly-scoped issues garner agreement that a concern is legitimate, these concerns are often nullified using the usual logics of "customer centricity", as in Section 3.2 when inclusive VR was dismissed as something the customer did not demand, or when management or clients dismissed the two technical solutions proposed to remedy concerns as requiring too many resources to implement, as in Section. 4.1. In these cases, incentives won out, even for concerns aimed at improving the product rather than critiquing its entirety.

Therefore, fixing "ethics bugs" often relies on practitioners' *power* to persuade others to dedicate resources to fixing them, and this in turn motivates further work to develop tactics of persuasion such as justifying solutions to ethics problems in terms of organizational incentives (see Sec. 4.2, see also [71]), and work to quantify and provide outside evidence for relationship between ethics fixes (*i.e.*, accessibility, see Sec. 3.2) and the incentives that decision makers care about, such as product success or user growth. Our work also helps answer calls for "guidance around how to navigate organizational power dynamics" [72] when raising ethical concerns that toolkits help identify, by helping understand the power structures into which ethics interventions must work within, and the limits on the power of those who may apply them. Additionally, practitioners may themselves have power to prioritize among bugs [25], and our work suggests the opportunity to examine where "ethics bugs" lie in their prioritization. Others show that practitioners advocate for ethics less powerfully due to career concerns [42, 56], we show this is inflected by financial and immigration precarity (see Sec. 5.1, also [11]) and workplace culture (see Sec. 5.2). This suggests future research should investigate other contingencies on practitioner's power to advocate for ethics.

This also has implications for education. A recent survey of undergraduate tech ethics courses found their "overarching goal [...] appears to be to teach students to recognize ethical issues in the world" [18], but fewer than one quarter touch on the systems of power – "capitalism, financial models, marketing, pricing" – within which issues must be addressed. Our study enumerates ethical concerns that practicing software engineers face in Section 3.1, which can help ensure in class examples are representative of the concerns that practicing software engineers face at work. However, lest courses help students identify issues but leave them unprepared to advocate for fixes, the factors we enumerate in Section 5 can help tech ethics teach students how they may encounter these systems of power in their future careers, alongside learning tactical skills to raise concerns, which we detail in Sections 4.1,4.2, and 4.3.

6.2 Labor as counterpower to question an industry's *raison d'être*

Practitioners also raised larger concerns which question the *raison d'être* of their organization or industry. Other scholars have critiqued design-stage interventions as insufficient [19], especially when harm is inherent in how systems are used [70], aligned with calls for ethics work to "move away from prioritizing notions of good design" and towards critique of "what and whose goals are being achieved" [51]. Concerns we collect at this end of the spectrum provide myriad examples of practitioners raising these critiques. Even with the above ways to improve tech ethics interventions, the kinds of ethics concerns addressable using them are likely to remain limited to those aligning with the company's incentives. Therefore, additional research and education is needed to account for ethical concerns which may threaten a company's *raison d'être*.

Our empirical evidence demonstrates that when practitioners develop concerns with their company or industry's business practices, they see few options other than withholding their labor (*i.e.*, resigning and finding a new job, see Sec. 4.3). Though this made some feel less culpable in harm, some believed they would be easily replaced and the system still built. Indeed, Palantir CEO Alex Karp said "I've had some of my favorite employees leave" over the

company's contract to provide software to US Immigration and Custom's Enforcement that helped separate immigrant children from their families [2, 7], but the contract continued. However, we hear from other participants that resigning can be powerfully disruptive, leading some clients to cancel precarious projects (see Sec. 4.3). Future research should explore individuals' power over outcomes: when is a resignation by a concerned engineer successful in halting a software project? This can build on research examining volunteer open source projects, which studies how the departure of crucial "truck factor" developers often puts the project into serious peril [4]. Some of our participants practice "feet voting" by proactively planning their career in alignment with their values (see Sec. 4.4), future work can evaluate how commonly and with what priority ethical concerns factor into tech job seekers' priorities, and examine support and information needs for ethically-concerned job seekers.

Practitioners' concerns with their industry's *raison d'être* also has implications for education. Very few of the 115 tech ethics courses surveyed in one study encouraged "students to create their own personal code of ethics" [18]. Given that our work shows that practitioners see their employment choices as opportunities to exercise agency in accordance with their ethical views, tech ethics courses may consider providing help with career planning as a primary opportunity to align their labor with their values (see Section 4.4), in addition to teaching about ethics-focused tech worker rights organizations such as the Tech Workers Coalition [1]. Tech education may also expand to teach skills identified in our study, such as negotiating for ethics using organizational incentives (see 4.2), but more powerfully, it can also call attention to strategies for building collective power, including watercooler talk to socialize concerns (see 4.6), whistle blowing and legal remedies (see 4.5) to discussion of tech worker unions (see 4.4).

6.3 The coherence of a focus on "AI" or "Big Tech" in tech ethics discourse

The power of labor is strongest when acting collectively: as one of our participants recognized, "the work doesn't get done without us" (see 4.4). However, only two of our participants raised unions as an avenue to advocate for ethics concerns (see Sec. 4.4) despite high-profile efforts to collectively organize over ethics issues at large firms such as Google or Microsoft [37]. Our study shows show that tech ethics research ought to: firstly, broaden to consider tech ethics beyond its contemporary focus on AI; and secondly, broaden beyond studying software engineers at "Big Tech" companies. This larger focus will examine more contingencies in tech worker power and enable a broader coalition by finding issues of common concern, but also shift consideration of ethics towards harm irrespective of implementation, instead of a privileged focus on "AI" concerns.

Firstly, to capture as wide of a scope of ethical concerns as possible, and given divergent conceptions of what "AI" is [36], we did not limit our study to "AI" practitioners, or to concerns related to "Ethical AI". While one analysis concluded that "activism" by "the artificial intelligence (AI) community" was "successful" in part because of "a coherent shared culture" bourne of attending the same conferences, and concluded that "The AI community is acting together – it is organised" [5], we argue this casts the AI community as a monolith, characterized by its most privileged academic members, and the sources of power and concerns they hold. Despite not deliberately recruiting AI practitioners, most of our interviewees were "building 'smart' machines" in some way (*i.e.*, as per [15]), and some positioned themselves as working on "AI" systems. Despite this, none of our "AI" participants consider themselves "organized" nor talked about themselves as part of a wider shared "AI" culture.

Given this, we argue that a focus on "AI" in tech ethics discourse implies a limited scope of scrutiny, focusing on design-stage interventions [19]. Using "AI" is a design choice, and whereas many of the concerns our participants raise do not depend on whether the system in question uses "AI" or not, especially when concerned with the *raison d'être* of their industry (see Sec. 3.2). Therefore, future work on the ethics with software practitioners should avoid limiting recruitment to AI practitioners or framing questions to exclusively AI concerns, as such a limitation may be artificial and limiting in the same way that AI principles may limit scrutiny to system design [24]. The scopes we present in Section 3.2 may help conceptualize practitioner concerns, beyond AI. Similarly, "AI Ethics" [9] courses may consider expanding to study tech ethics broadly, as some already do [18].

Secondly, only one of our interviewees currently works for a "Big Tech" company (*i.e.*, [59]), though he did not speak of concerns working there, and only one other spoke of concerns from past experience working in Big Tech. The majority of our interviewees were contingent contractors, working in a variety of B2B companies, or working as software engineers at non-tech companies (see Sec. 8.3). This is relevant in light of calls to do research beyond "large, internal software development teams" [63], but also given that many ethics issues "are important but arcane and not conducive to media coverage [...] in particular for low-visibility AI companies, including those that do not market to the public but instead sell their AI to governments or other companies." [14]. Major companies invest heavily in certain framings of AI ethics to the point they raise concerns of *capture* of not only AI resources [69] but also AI ethics discourse [74], and they also are the site of the most high-profile examples of countervailing collective organizing [3, 13, 37], and thus their workers may well be aware of (certain versions of) broader ethics discussions. Therefore, we argue that studies of practitioner ethics challenges, which often focus on "large U.S.-based technology companies" [71] or "major companies" [56] risk assuming a base level of exposure to AI ethics discourse, and thus risk assuming a certain level of generality around what ethics concerns exist.

Given this, we suggests that AI Ethics research may need to broaden to better account for the majority of software practitioners who do not work at "major" companies. For example, we believe that our participants' feelings of being isolated in their ethics concerns and resulting mental health consequences (see Sec. 4.6) and attempts to build this community by socializing their concerns (see Sec. 4.6) may reflect unique isolation in contrast to in tech-centric companies, where processing concerns with similarly aware colleagues may help [62]. To account for this, and to find ways to build collective power across diverse experiences, future work on software practitioners' ethics concerns ought to deliberately recruit from beyond tech companies, perhaps using existing catalogs of

collective actions from a broad variety of tech workers including blue and white collar tech workers [49, 50].

7 CONCLUSION

In this study, we report on the ethical concerns software engineers identify themselves, without the use of ethics interventions such as fairness checklists [42], codified principles [33], or institutionalized ethics programs [44], which others argue impose a limited scope of ethical scrutiny [24, 35]. Our results show that with an open ended scope, practitioners raise a wide variety of ethical concerns, including those which question the *raison d'être* of their company or industry. We examine the strategies practitioners use to seek to resolve their concerns, and the way in which personal precarity, workplace culture, and organizational incentives limit their power to do so. In our discussion, we highlight the centrality of *power*: our results suggest that ethics interventions, research, and education must expand from helping practitioners merely *identify* issues to instead helping them build their (collective) *power* to resolve them.

REFERENCES

[1] 2020. A Tech Workers' Bill of Rights. https://techworkerscoalition.org/bill-of-rights/. *Tech Workers Coalition* (2020). Accessed: 2021-010-1.

[2] Mike Allen. 2020. Palantir's CEO Said He's Suffered Because of His Contract with ICE. https://www.axios.com/2020/05/26/palantir-ceo-ice-immigration.

[3] Anat Alon-Beck. 2020. Times They Are a-Changin': When Tech Employees Revolt! *Md. L. Rev.* 80 (2020), 120.

[4] Guilherme Avelino, Leonardo Passos, Andre Hora, and Marco Tulio Valente. 2016. A novel approach for estimating truck factors. In *2016 IEEE 24th International Conference on Program Comprehension (ICPC)*. IEEE, 1–10.

[5] Haydn Belfield. 2020. Activism by the AI Community: Analysing Recent Achievements and Future Prospects. In *AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York NY USA, 15–21. https://doi.org/10.1145/3375627.3375814

[6] Garfield Benjamin. 2022. #FuckTheAlgorithm: Algorithmic Imaginaries and Political Resistance. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Seoul Republic of Korea, 46–57. https://doi.org/10.1145/3531146.3533072

[7] Sam Biddle and Ryan Devereaux. 2019. Peter Thiel's Palantir Used to Bust Migrant Children's Relatives. https://theintercept.com/2019/05/02/peter-thiels-palantir-was-used-to-bust-hundreds-of-relatives-of-migrant-children-new-documents-show/.

[8] Abeba Birhane, Elayne Ruane, Thomas Laurent, Matthew S. Brown, Johnathan Flowers, Anthony Ventresque, and Christopher L. Dancy. 2022. The Forgotten Margins of AI Ethics. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Seoul Republic of Korea, 948–958. https://doi.org/10.1145/3531146.3533157

[9] Jason Borenstein and Ayanna Howard. 2021. Emerging Challenges in AI and the Need for AI Ethics Education. *AI and Ethics* 1, 1 (Feb. 2021), 61–65. https://doi.org/10.1007/s43681-020-00002-7

[10] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. (2012).

[11] Enda Brophy. 2006. System Error: Labour Precarity and Collective Organizing at Microsoft. *Canadian Journal of Communication* 31, 3 (Oct. 2006), 619–638. https://doi.org/10.22230/cjc.2006v31n3a1767

[12] Margaret Burnett, Simone Stumpf, Jamie Macbeth, Stephann Makri, Laura Beckwith, Irwin Kwan, Anicia Peters, and William Jernigan. 2016. GenderMag: A Method for Evaluating Software's Gender Inclusiveness. *Interacting with Computers* 28, 6 (Nov. 2016), 760–787. https://doi.org/10.1093/iwc/iwv046

[13] Kelley Changfong-Hagen. 2020. " Don't Be Evil": Collective Action and Employee Prosocial Activism. *HRLR Online* 5 (2020), 188.

[14] Peter Cihon, Jonas Schuett, and Seth D. Baum. 2021. Corporate Governance of Artificial Intelligence in the Public Interest. *Information* 12, 7 (July 2021), 275. https://doi.org/10.3390/info12070275

[15] Jenny L Davis. 2020. *How artifacts afford: The power and politics of everyday things*. MIT Press.

[16] Julia Dressel and Hany Farid. 2018. The Accuracy, Fairness, and Limits of Predicting Recidivism. *Science Advances* 4, 1 (Jan. 2018), eaao5580. https://doi.org/10.1126/sciadv.aao5580

[17] Avi Feller, Emma Pierson, Sam Corbett-Davies, and Sharad Goel. 2016. A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. *The Washington Post* 17 (2016).

[18] Casey Fiesler, Natalie Garrett, and Nathan Beard. 2020. What Do We Teach When We Teach Tech Ethics?: A Syllabi Analysis. In *51st ACM Technical Symposium on Computer Science Education*. ACM, Portland OR USA, 289–295. https://doi.org/10.1145/3328778.3366825

[19] Ben Gansky and Sean McDonald. 2022. CounterFAccTual: How FAccT undermines its organizing principles. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. 1982–1992.

[20] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. Datasheets for datasets. *Commun. ACM* 64, 12 (2021), 86–92.

[21] Marcela F González. 2022. Precarity for the global talent: The impact of visa policies on high-skilled immigrants' work in the United States. *International Migration* 60, 2 (2022), 193–207.

[22] DW Gotterbarn, Bo Brinkman, Catherine Flick, Michael S Kirkpatrick, Keith Miller, Kate Vazansky, and Marty J Wolf. 2018. ACM code of ethics and professional conduct. (2018).

[23] Colin M Gray, Yubo Kou, Bryan Battles, Joseph Hoggatt, and Austin L Toombs. 2018. The dark (patterns) side of UX design. In *2018 CHI conference on human factors in computing systems*. 1–14.

[24] Daniel Greene, Anna Lauren Hoffmann, and Luke Stark. 2019. Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. In *52nd Hawaii international conference on system sciences*.

[25] Philip J Guo, Thomas Zimmermann, Nachiappan Nagappan, and Brendan Murphy. 2010. Characterizing and predicting which bugs get fixed: an empirical study of microsoft windows. In *32Nd ACM/IEEE International Conference on Software Engineering-Volume 1*. 495–504.

[26] Karen Hao. 2019. In 2020, let's stop AI ethics-washing and actually do something. *MIT Technology Review* 27, December (2019), 2019.

[27] Thomas A. Hemphill. 2019. 'Techlash', Responsible Innovation, and the Self-Regulatory Organization. *Journal of Responsible Innovation* 6, 2 (May 2019), 240–247. https://doi.org/10.1080/23299460.2019.1602817

[28] Ralph Hertwig, Carola Fanselow, and Ulrich Hoffrage. 2003. Hindsight bias: How knowledge and heuristics affect our reconstruction of the past. *Memory* (2003).

[29] Claudia Hilderbrand, Christopher Perdriau, Lara Letaw, Jillian Emard, Zoe Steine-Hanson, Margaret Burnett, and Anita Sarma. 2020. Engineering Gender-Inclusivity into Software: Ten Teams' Tales from the Trenches. In *ACM/IEEE 42nd International Conference on Software Engineering*. ACM, Seoul South Korea, 433–444. https://doi.org/10.1145/3377811.3380371

[30] Kenneth Holstein, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudik, and Hanna Wallach. 2019. Improving fairness in machine learning systems: What do industry practitioners need?. In *2019 CHI conference on human factors in computing systems*. 1–16.

[31] Lilly Irani. 2019. *Chasing Innovation: Making Entrepreneurial Citizens in Modern India*. Princeton University Press. https://doi.org/10.1515/9780691189444

[32] Maurice Jakesch, Zana Buçinca, Saleema Amershi, and Alexandra Olteanu. 2022. How Different Groups Prioritize Ethical Values for Responsible AI. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 310–323. https://doi.org/10.1145/3531146.3533097

[33] Anna Jobin, Marcello Ienca, and Effy Vayena. 2019. The global landscape of AI ethics guidelines. *Nature Machine Intelligence* 1, 9 (2019), 389–399.

[34] Jason Karaian and Lora Kelley. 2023. How Big Tech Layoffs Stack Up With the Rest of Their Work Forces. https://www.nytimes.com/2023/01/21/business/tech-layoffs.html.

[35] Os Keyes, Jevan Hutson, and Meredith Durbin. 2019. A mulching proposal: Analysing and improving an algorithmic system for turning the elderly into high-nutrient slurry. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–11.

[36] P. M. Krafft, Meg Young, Michael Katell, Karen Huang, and Ghislain Bugingo. 2020. Defining AI in Policy versus Practice. In *AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York NY USA, 72–78. https://doi.org/10.1145/3375627.3375835

[37] Logan Kugler. 2021. The unionization of technology companies. *Commun. ACM* 64, 8 (2021), 18–20.

[38] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *33rd annual ACM conference on human factors in computing systems*. 1603–1612.

[39] Michelle Seng Ah Lee and Jat Singh. 2021. The landscape and gaps in open source fairness toolkits. In *2021 CHI conference on human factors in computing systems*. 1–13.

[40] Zachary C Lipton. 2018. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* 16, 3 (2018), 31–57.

[41] Michael Madaio, Lisa Egede, Hariharan Subramonyam, Jennifer Wortman Vaughan, and Hanna Wallach. 2022. Assessing the Fairness of AI Systems: AI Practitioners' Processes, Challenges, and Needs for Support. *ACM Conference on Human-Computer Interaction* 6, CSCW1 (2022), 1–26.

[42] Michael A Madaio, Luke Stark, Jennifer Wortman Vaughan, and Hanna Wallach. 2020. Co-designing checklists to understand organizational challenges and opportunities around fairness in ai. In *2020 CHI Conference on Human Factors in Computing Systems*. 1–14.
[43] Andrew McNamara, Justin Smith, and Emerson Murphy-Hill. 2018. Does ACM's code of ethics change ethical decision making in software development?. In *2018 26th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering*. 729–733.
[44] Jacob Metcalf, Emanuel Moss, et al. 2019. Owning ethics: Corporate logics, silicon valley, and the institutionalization of ethics. *Social Research: An International Quarterly* 86, 2 (2019), 449–476.
[45] Debra E Meyerson. 2008. *Rocking the boat: How tempered radicals effect change without making trouble*. Harvard Business Review Press.
[46] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model cards for model reporting. In *conference on fairness, accountability, and transparency*. 220–229.
[47] Brent Mittelstadt. 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence* 1, 11 (2019), 501–507.
[48] Anton J Nederhof. 1985. Methods of coping with social desirability bias: A review. *European journal of social psychology* (1985).
[49] Nataliya Nedzhvetskaya, JS Tan, Hyatt Dirbas, and Wynnie Chan. 2022. Collective Action in Tech. https://data.collectiveaction.tech/.
[50] Nataliya Nedzhvetskaya and J. S. Tan. 2022. The Role of Workers in AI Ethics and Governance. In *The Oxford Handbook of AI Governance* (first ed.), Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young, and Baobao Zhang (Eds.). Oxford University Press, C68.S1–C68.N14. https://doi.org/10.1093/oxfordhb/9780197579329.013.68
[51] Gina Neff. 2020. From Bad Users and Failed Uses to Responsible Technologies: A Call to Expand the AI Ethics Toolkit. In *AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New York NY USA, 5–6. https://doi.org/10.1145/3375627.3377141
[52] Rodrigo Ochigame. 2019. How Big Tech Manipulates Academia to Avoid Regulation. https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/.
[53] Will Orr and Jenny L Davis. 2020. Attributions of ethical responsibility by Artificial Intelligence practitioners. *Information, Communication & Society* 23, 5 (2020), 719–735.
[54] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2022. Designing Fair AI in Human Resource Management: Understanding Tensions Surrounding Algorithmic Evaluation and Envisioning Stakeholder-Centered Solutions. In *CHI Conference on Human Factors in Computing Systems*. 1–22.
[55] Hatim A. Rahman. 2021. The Invisible Cage: Workers' Reactivity to Opaque Algorithmic Evaluations. *Administrative Science Quarterly* 66, 4 (Dec. 2021), 945–988. https://doi.org/10.1177/00018392211010118
[56] Bogdana Rakova, Jingying Yang, Henriette Cramer, and Rumman Chowdhury. 2020. Where Responsible AI meets Reality: Practitioner Perspectives on Enablers for shifting Organizational Practices. *In 24th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (2020).
[57] Brianna Richardson, Jean Garcia-Gathright, Samuel F. Way, Jennifer Thom, and Henriette Cramer. 2021. Towards Fairness in Practice: A Practitioner-Oriented Rubric for Evaluating Fair ML Toolkits. In *2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–13. https://doi.org/10.1145/3411764.3445604
[58] Marc J Riemer. 2007. Communication skills for the 21st century engineer. *Global J. of Engng. Educ* 11, 1 (2007), 89–100.
[59] Henrik Skaug Sætra, Mark Coeckelbergh, and John Danaher. 2022. The AI Ethicist's Dilemma: Fighting Big Tech by Supporting Big Tech. *AI and Ethics* 2, 1 (Feb. 2022), 15–27. https://doi.org/10.1007/s43681-021-00123-7
[60] Niloufar Salehi, Lilly C Irani, Michael S Bernstein, Ali Alkhatib, Eva Ogbe, and Kristy Milland. 2015. We are dynamo: Overcoming stalling and friction in collective action for crowd workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1621–1630.
[61] Reginald G Smart. 1966. Subject selection bias in psychological research. *Canadian Psychologist/Psychologie canadienne* (1966).
[62] Norman Makoto Su, Amanda Lazar, and Lilly Irani. 2021. Critical Affects: Tech Work Emotions Amidst the Techlash. *ACM Conference on Human-Computer Interaction* 5, CSCW1 (2021), 1–27.
[63] Suzanne L Thomas. 2019. Migration versus management: the global distribution of computer vision engineering work. In *2019 ACM/IEEE 14th International Conference on Global Software Engineering (ICGSE)*. IEEE, 12–17.
[64] Ville Vakkuri, Kai-Kristian Kemell, Marianna Jantunen, and Pekka Abrahamsson. 2020. "This is Just a Prototype": How Ethics Are Ignored in Software Startup-Like Environments. In *Agile Processes in Software Engineering and Extreme Programming*, Viktoria Stray, Rashina Hoda, Maria Paasivaara, and Philippe Kruchten (Eds.). Vol. 383. Springer International Publishing, Cham, 195–210. https://doi.org/10.1007/978-3-030-49392-9_13 Series Title: Lecture Notes in Business Information Processing.
[65] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and accountability design needs for algorithmic support in high-stakes public sector decision-making. In *2018 chi conference on human factors in computing systems*. 1–14.
[66] Ben Wagner. 2018. Ethics as an escape from regulation. From "ethics-washing" to ethics-shopping? (2018).
[67] Robert Philip Weber. 1990. *Basic content analysis*. Number 49. Sage.
[68] Robert S Weiss. 1995. *Learning from strangers: The art and method of qualitative interview studies*. Simon and Schuster.
[69] Meredith Whittaker. 2021. The steep cost of capture. *Interactions* 28, 6 (2021), 50–55.
[70] David Gray Widder, Dawn Nafus, Laura Dabbish, and James Herbsleb. 2022. Limits and Possibilities for "Ethical AI" in Open Source: A Study of Deepfakes. In *conference on fairness, accountability, and transparency*.
[71] Richmond Y Wong. 2021. Tactics of Soft Resistance in User Experience Professionals' Values Work. *ACM Conference on Human-Computer Interaction* 5, CSCW2 (2021), 1–28.
[72] Richmond Y. Wong, Michael A. Madaio, and Nick Merrill. 2023. Seeing Like a Toolkit: How Toolkits Envision the Work of AI Ethics. *ACM Conference on Human-Computer Interaction* 7, CSCW1 (April 2023), 1–27. https://doi.org/10.1145/3579621
[73] Karen Yeung, Andrew Howes, and Ganna Pogrebna. 2019. AI governance by human rights-centred design, deliberation and oversight: An end to ethics washing. *The Oxford Handbook of AI Ethics, Oxford University Press (2019)* (2019).
[74] Meg Young, Michael Katell, and P.M. Krafft. 2022. Confronting Power and Corporate Capture at the FAccT Conference. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Seoul Republic of Korea, 1375–1386. https://doi.org/10.1145/3531146.3533194
[75] Thomas Zimmermann. 2016. Card-sorting: From text to themes. In *Perspectives on data science for software engineering*. Elsevier, 137–141.
[76] J Zunger. 2018. Computer science faces an ethics crisis. The Cambridge Analytica scandal proves it. *Boston Globe* 22 (2018).

8 APPENDIX

8.1 Survey Questions

(1) Have you ever had ethical concerns about a software system you were asked to contribute to?
(2) What were you asked to do?
(3) What were your concerns?
(4) What action, if any, did you take as a result of your concerns?
(5) How, if at all, were your concerns resolved?
(6) How did you feel about this outcome?
(7) Was there anything that made it feel *easier* to act on your concerns?
(8) Was there anything that made it feel *harder* to act on your concerns?
(9) Which best describes your employment status?
(10) If employed, which best describes the industry you work in?
(11) If employed, approximately how many people work for your employer?
(12) Including any education, for how many years have you been coding?
(13) Which continent do you live in?
(14) Which best describes you? [Gender]
(15) If you wish, feel free to explain any of your above answers.

8.2 Interview Guide

These questions were used as starting points for a semi-structured interview [68], with many additional prompts as the interview progressed.

(1) To start off, can you tell me a bit about yourself, and about your background?
(2) Let's talk about the experience you wrote about on your survey. You said that you were asked to do [summary of

task]. First, can you set the scene and tell me a little about the circumstances of your employment?
(3) On the survey, you wrote that you were concerned about [brief 3-5 word summary of concern(s)]. Can you walk me through how this concern first arose?
(4) Do you think there was anything you could have done, within the purview of your assigned responsibilities, to resolve your ethical concerns? Or was it more or less out of your hands?
(5) Stepping back, why did you think this was an issue? How did you come to think of this as an ethical concern?
(6) In the survey you wrote about [factor] making it *easier* to act. Can you talk a little more about that? Was there anything else that made it *easier* to act?
(7) You also wrote about [factor] making it *harder* to act. Can you talk a little more about that? Was there anything else that made it *harder* to act?
(8) Do you have any friends or colleagues who have been in a similar situation?

8.3 Interview Participant Demographics

Gender	Highest Degree	Seniority	Sector	Role	Yrs Coding	Org. Size	Concern(s)
Male	MS	Sr.	Government	ML Researcher	10	100-499	inequality, surveillance
Male	HS	Sr.	Government	CTO	35	20-99	surveillance
Male	BS	Sr.	Government	CTO	20	100-499	legal
Male	BS	Mid.	Government	Software Eng.	8	10,000+	security
Male	BA	Jr.	Military	Software Eng.	20	10,000+	military
Male	BS	Jr.	Military	Software Eng.	6	1,000-4,999	military
Male	PhD	Sr.	Edtech	CTO	37	<10	privacy
Female	MS	Jr.	Edtech	VR Developer	14	1,000-4,999	accessibility, inclusivity
Male	MS	Sr.	Academia	Researcher	17	500-999	surveillance
Male	BS	Jr.	Academia	Researcher	8	10,000+	research ethics
Male	BS	Jr.	Insurance	Software Consult.	22	100-499	insurance denial
(Declined)	MS	Jr.	Fintech	Data Scientist	18	10,000+	inequality
N.B. (femme)	MS	Mid.	Banking	Data Scientist	12	1,000-4,999	inequality
Male	BS	Sr.	Humanitarian	Software Eng.	10	1,000-4,999	labor exploitation
Nonbinary	BA	Mid.	Health nonprofit	Software Config.	6	10,000+	life safety
Nonbinary	BS	Jr.	Security	Software Eng.	9	10,000+	privacy, labor
Male	HS	Mid.	Construction	Software Eng.	15	10-19	privacy
Male	PhD	Sr.	Mobile dev.	Data Scientist	25	100-499	privacy
Male	BS	Jr.	Networking	Software Eng.	12	500-999	privacy
Male	BS	Jr.	Video software	Software Eng.	6	20-99	manipulation, misuse
Male	HS	Mid.	Agriculture	Software Eng.	7	<10	environment, labor exploitation

Table 1: Interview Participant demographics grouped by sector. To protect anonymity, we do not provide participant numbers nor uniquely identify their continents (spanning Africa, Australia, Europe, with the majority in North America) in this table.