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

ANONYMIZED FOR PEER REVIEW

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 ability 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 power to resolve them.

1 INTRODUCTION

In light of public pressure and negative press [35], many large technology companies are attempting to address harms from algorithmic systems, such as by instituting ethics initiatives which converge on principles such as transparency or fairness [42]. However, some argue that the convergence around principles obscures underlying political and normative disagreements [60], that principles may have discordant definitions [20, 21, 50], or that principles limit scrutiny to a system's design and thus avoid scrutiny on what business use the systems for [31], or who they sell them to. Others point out that fair or transparent systems may nonetheless be put to unethical *uses* [43, 88]. These concerns lead to accusations of "ethics washing" [34, 81]: where companies put forward voluntary codes to burnish their reputation and avoid regulation [65], without changing behavior [93].

Additionally, there is evidence that even "major companies" sometimes see ethics work as "too complicated for the organization's current level of resources" despite having official ethics initiatives [69], and that ethics questions may be seen as premature in startup environments [78].

Given that software practitioners are able to directly affect how systems operate as they are the ones building them, research has examined the needs of these practitioners as they attempt to put ethical principles like fairness into practice [39, 51, 52, 69], and others study the effect of ethical codes on practitioners' ethical decision making [54]. However, principles or codes may impose a narrow scope and thus foreclose on wider concerns [31], and many software practitioners work at smaller companies that may not have official ethics initiatives. Our paper responds to these issues by studying the kinds of ethics concerns practitioners identify on their own in the course of their work, what they do

Author's address: Anonymized for peer review.

about these concerns, and what factors affect their ability to resolve them. Accordingly, we ask the following research questions:

RQ1: What do software engineers see as an ethical concern?

RQ2: What happens when software engineers develop ethical concerns?

RQ3: What affects software engineers' power to resolve their concerns?

We are not aware of previous studies that take this open scope to examining software practitioners' ethical concerns. A better understanding of practitioner ethical concerns will identify ways to expand or enhance current ethics supports, such as ethics checklists [52], toolkits [92], or education [22] and identify concerns which may require more powerful strategies [46, 91]. Also, by focusing on non-hypothetical cases where practitioners become concerned that a system they work on may be causing harm, we can examine challenges based on experience, instead of those prompted by fairness checklists or other tools, which often seek to anticipate harm in the abstract or in advance.

Our paper begins by reviewing related work in Section 2, and outlining our methods to survey and interview software engineers in Section 3. We then explain our findings for each of our research questions. In Section 4, we explain the kinds of concerns our participants hold, and how they vary from smaller bugs to wider critiques of their industry. In Section 5, we explain the broad range of actions that practitioners take in response to their concerns, from proposing technical solutions, to negotiating their concerns with managers and clients, to quitting and finding a job that aligns with their values. In Section 6, we explain how factors such as personal precarity, workplace culture, and organizational incentive structures affect practitioners' ability to raise and resolve their concerns. In Section 7, we discuss how our results show that interventions to identify ethics issues are insufficient if practitioners lack authority to dedicate resources to fixing them, and that withholding one's labor is often the only recourse to many practitioners' larger concerns. We finally discuss how our results show that research on practitioners' ethics concerns ought to expand beyond centering those at major technology companies, and to concerns broader than "AI" as this may direct critique artificially to system design rather than use.

2 RELATED WORK

Langon Winner argued that technological artifacts embody "politics": their design makes certain political and ethical outcomes more or less likely [90]. One example of this could be that facial recognition systems have disparately high error rates for female or darker-skinned faces [13], another shown in studies of how software may embed gender biases in its interface [14]. In response, within the field of HCI, the study of value sensitive design is one way to align a system's values with those of the stakeholders it affects [24, 79], in light of possible conflicts. For example, surveys by Jakeech *et al.* showed that AI practitioners' have different values than of the general public in AI system design [41], and a recent study showed that in the context of algorthmic evaluation of worker performance, workers have different concerns than those who seek to monitor them [67].

Some have studied difficulties encountered in tech company programs to translate ethical principles into practice in software companies [69], and the incentives which complicate the work of those trying to do so [55]. There is also work on similar challenges in software practitioner's day to day work, who may have some power over well studied ethical concerns such as in manipulative "dark patterns" in UX design [29, 33]. For example, the Association for Computing Machinery publishes an ethical code [28], though one study found it has limited effect on software engineer's choices in hypothetical ethical scenarios [54]. Veale *et al.* found that public sector AI practitioners have challenges embedding values in their work [80] in ways that needed researcher attention. Holstein *et al.* found practitioners have unmet needs when trying to build fair AI systems, such as help curating high quality data and integrating subject

matter expertise, when building fair AI systems [39]. Others narrow in on challenges employing specific fairness methods like disaggregated evaluations [51] Madaio *et al.* co-designed a fairness checklist with practitioners [52], finding checklists to empower individual advocates, but found concerns that using them may be seen as a guarantee for fairness. Su *et al.* study how public rage about tech harms ("techlash") affects tech workers. Richmond Wong identified UX professional's tactics of "soft resistance" (*i.e.*, [61]), where they seek to subvert practices for ethical ends. There have been a variety of ways to help practitioners identify and document ethical issues in AI systems and data, such as dataset datasheets [26] and model cards [59], and ethics toolkits [92]. Other work asks engineers how they attribute ethical responsibility for AI systems they work on, but not with reference to a specific ethical concern they hold themselves [66], and other work shows how codified norms, such as those inscribed in open source licensing, lead practitioners to disavow responsibility for downstream harms [88]. There are also increasingly studies of trust in AI systems in software engineering workplaces [82, 86].

There has been analyses of activism within the AI community [7], but also study of concerns and activism from less powerful tech workers. Gig workers find algorithmic management opaque and an "invisible cage" [68], other work identifies how they used online forums to share information [47]. Salehi *et al.* study challenges to collective action for online crowd workers [73], and Nedzhvetskaya *et al.* collect examples of tech worker collective actions, both in blue and white collar contexts [63]. Our work attempts to answer recent calls from within HCI to support all kinds of tech labor through worker centered design [23], and a series of calls for HCI research to consider social impacts and responsibility in technological innovation [5, 44, 89].

3 METHODS AND PARTICIPANTS

Under an interpretivist epistemological paradigm [49], we conducted a survey to reach a broad range of software engineers concerns they face, with an optional follow-up interview to collect more narrative detail and probe on their response to their concerns, and their feelings and actions since their concern. We recruited participants for our survey using multiple recruitment 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 from practioners with ethical concerns. They survey asked about respondents' concerns, actions they took, resolutions, and demographics (instrument in Appendix 8.1). 90 were employed full time, 8 part time, 7 contractors, 6 students, 2 retired, 1 looking for work, 1 did not respond. Participants were employed at companies of many different sizes: 7 at a company with fewer than 10 employees, 6 at 10-19, 23 at 20-99, 20 at 100-499, 11 at 500 to 999, 9 at 1000 to 4999, 6 at 5000 to 9999, 20 at a firm with more than 10,000 employees. Participants were relatively experienced, reporting a mean of 17 years of experience coding (median = 15, minimum = 4, maximum = 46 years). Respondents spanned six continents: 68 participants lived in North America, 34 in Europe, 4 in Australia, 4 in Asia, 4 in South America, and 2 in Africa. 80 participants identified as male, 10 as female, 6 as nonbinary or gender nonconforming, and 5 self-described. 43 survey respondents signed up for an optional follow-up interview, of which 21 responded to our follow-up request and completed an interview (see interviewee demographics in Table 1). We conducted semi-structured teleconference interviews [84] to collect a more detailed order of events as practitioners navigated their concerns, and to probe into the participants recollections of their internal state, and 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 and median: 41 minutes).

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	CS Researcher	17	500-999	surveillance
Male	BS	Jr.	Academia	PhD Student	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 the 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.

We analyzed survey and interview responses sequentially. We performed an *open qualitative card sort* on survey responses, an interpretive approach to categorize and relate qualitative data. [6, 38, 58, 87]. The two leading authors collaboratively sorted the respondents' survey answers into categories in a collaborative process; discussing and negotiating disagreements, and adjusting categories and resorting cards as necessary. This yielded finer grained categories under our research questions: categories of ethical concern (described in Table. 2), actions taken (used in Sec. 5), factors affecting perceived power (Sec. 6), and any resolutions (Sec. 5). For the interview transcripts, the two leading authors jointly performed three rounds of iterative [83], thematic analysis on this data [11, 75]. The first round comprised of *open coding*: annotating meaningful segments of data using short annotations, including connections possible concepts of interests, including our three research questions (introduced above). In the second round, open codes were discussed and merged with themes identified in the survey card sort to arrive at a *closed coding frame*, then applied to code all transcripts using Qualitative Content Analysis software. In addition, the interviewer took notes during and after each interview, and all authors gathered at least weekly to discuss and iterate on the interview protocol. In the ensuing sections, quotes from the survey (S#) and interview (I#) are reported together.

With our interpretivist paradigm [49], our findings emerge from the intersubjectivity between researchers and participants, and cultural frames we do and do not share. Our study *makes use of* self-selection [74] to surface self-identified ethical concerns without any pre-ordained scope, but therefore our results *do not* support general claims about software engineers (*i.e.*, prevalence of a given concern). Methods such as interviews and surveys rely on self-reported experiences, risking social desirability (*i.e.*, representing past behavior as pro-social [62]), and hindsight biases (*i.e.*,

ascribing too much predictability to contingent past events [36]). Surveys and interviews were conducted in English, a widely-spoken language for intercultural communication in engineering [71], but we acknowledge that this may lead to our work inadequately representing the perspectives of software engineers not proficient in English or active in the English language venues where we recruit. We discuss the relative privilege of our participants in Section 7.3).

4 RQ1: WHAT DO SOFTWARE ENGINEERS SEE AS AN ETHICAL CONCERN?

Our respondents reported a wide range of ethical concerns, from software for automating weapons of war to security vulnerabilities resulting from using deprecated frameworks. We observe that "ethical concerns" about a system are inextricably linked to the idea of "harm" - hypothetical or definite negative repercussions resulting from its development or deployment. We categorized the most frequently enumerated concerns in Table 2 thematically by type of concern.

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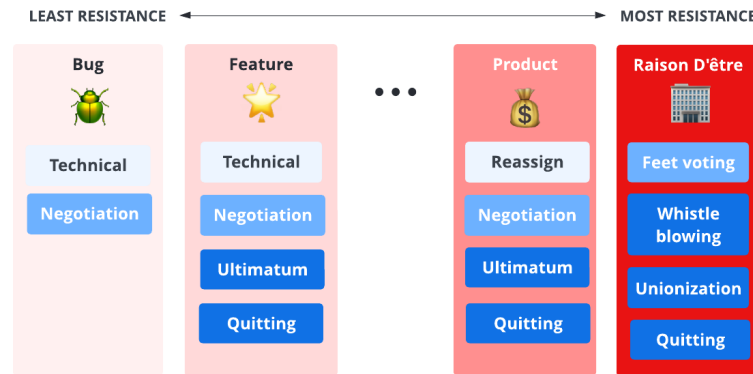

Fig. 1. Scopes of concerns, on a spectrum illustrating how resolvable the concern is within organizational incentives. Each box delineates a scope, coupled with which actions are sensible for each scope.

Our respondents performed software work within organizations, either as full time employees, independent contractors or as consultants. As a result, their work was often defined externally by management or clients, rather than internally by their own wishes or desires.

We found that ethical concerns varied wildly in their *scope* based on how much of what an organization stands for a given concern concern calls into question. We found scope to be important in affecting outcomes, since concerns that are the most difficult to resolve almost always call into question deep organizational norms and incentive structures (Sec. 6.3). The scope of ethical concerns also impacts what kinds of actions practitioners feel are reasonable (Sec. 5). We illustrate the gamut of scopes by looking at four varying scopes of ethical concerns: those arising from bugs, features, products and finally, practitioner's concerns with their organization's *raison d'être*.

4.1.1 Bugs. Ethical concerns about "bugs" are when a system is not behaving as intended, and can often be fixed by changes to how a system is implemented. For example, one interviewee had concerns about the malfunction of an industrial crane control software he worked on: *"I eventually discovered an issue in their data for the cranes that just wasn't logically consistent." (I10)* Bug related concerns can also extend to robustness and security, which affect the

Theme and Respondents	Definition of Concern Theme, Representative Survey Quotes
Military (S10), (S24), (S35), (S38), (S40), (S42), (S44), (S52), (S56), (S66), (S83), (S88), (S91), (S92), (S94), (S95), (105)	Concerns about military uses of products, supporting war, injuring civilians, and wider critiques the military. *"software in support of simulations used to train US warfighters" (S88)* *"contributing to weapons systems [for] wars that should have never been started to begin with is wrong" (S44)* *"[military companies] suck huge amounts of money from public health care system for a near worthless service." (S105)*
Privacy (S5), (S12), (S18), (S30), (S32), (S39), (S46), (S55), (S69), (S71), (S75), (S86), (S100), (S111)	Concerns about inappropriate collection, storage, or misuse of user information, consent or data theft. *"contractors [were] to label hundreds of thousands of [private] video clips" (S12)* *"force users to enter personal info [...] not necessary to have for the task at hand [...] for marketing purposes" (S32)*
Advertising (S1), (S9), (S16), (S23), (S26), (S34), (S41), (S43), (S49), (S60), (S61), (S79), (S104), (S113)	Concerns about ways manipulative or annoying ways to acquire customers or encourage purchases. *"bypass [spam] prevention measures" (S79)* *"push users to buy something because stock was "almost out" [when this was untrue]" (S1)*
Surveillance (S3), (S6), (S13), (S37), (S45), (S47), (S48), (S53), (S76), (S80), (S85)	Concerns about surveillance to used for algorithmic management or political repression. *"observing how well grocery stockers stayed on task" (S45)* *"measure employees' pee time" (S37)* *"tracks telecom user sessions [and] be able to keep tracking after a restart." (S4)*
Environment (S8), (S11), (S101), (S102)	Concerns about enabling environmental harm. *"monitoring system for agropecuary [livestock] business [which] is highly damaging to the environment" (S8)*
Labor Displacement (S28), (S65), (S97)	Concerns about displacing human workers. *"I thought the software system could very well put some people out of a job" (S97)*
Algorithmic Decision Making (S19), (S20), (S36), (S69)	Concerns about algorithms making high impact decisions, often in public sector settings. *"track and collate people's answers when applying for benefits [to] find reasons to deny them" (S36)*
Legality (S54), (S70)	Concerns about being asked to create illegal software. *"every single step was clearly illegal" (S54)*
Security (S17), (S81)	Concerns about security vulnerabilities. *"write insecure/vulnerable code so I could deliver a feature for a deadline" (S17)*
Academia (S50), (S68)	Concern about systems which may undermine research integrity. *"use a more complicated and newer-sounding model 'to more easily make it through the review process.' [...] I feel like I am contributing to elitism in ML research and that my advisor is a sham" (S50)*

Table 2. Categories of concern in from card sort of survey responses where there was enough detail to discern. Only categories with at least two responses are included.

likelihood of malfunction in the future: *"software that are out of support, obsolete or deprecated are vulnerable and prone to security risks."* (S81)

In most cases, practitioners are aligned with their organizations over fixing bugs, since maintaining a system or product's intended functionality is within an organization's best interests. In the crane example, the practitioner described how a shared commitment to engineering excellence lead to a successful resolution: *"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)

This pattern of work — finding and fixing bugs — may be so routine that many practitioners do not register bugs as ethical issues. However, two practitioners described fixing bugs 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).

However, in some cases organizational incentives and norms can instead stifle practitioners' efforts to identify, fix and prevent bugs. For instance, one respondent wrote about how 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 spoke about how thin margins at his consulting firm made it difficult to do work of acceptable quality, making it easy to oversell to clients.

4.1.2 A specific feature. Disagreements between practitioner and their organization about what counts as a harm, and what merits "fixing" breaks down more significantly at the feature level. Unlike bugs, features are intentional aspects or behaviors of a product, commissioned by clients or product managers for the broad purpose of advancing organizational incentives. Thus, practitioners who contest features are often more directly calling into question their organizations' objectives and red lines.

For example, one interviewee recounted being asked to build a feature for insurance adjustment software that would round down GPS coordinates on properties being evaluated for insurance eligibility, which at certain edge cases *"would have denied people access to certain types of insurance."* (I10). Another interviewee working on workplace compliance software reported being asked implement a feature that he felt was invasive to employee privacy: *"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, log into a work account on the app."* (I6). What makes these two concerns "features" rather than "bugs" is intentionality — practitioners were deliberately asked to build out aspects of the product they felt were harmful.

Feature level concerns can also arise from practitioners' desire to embed values into the systems they work on. One interviewee developed ethical concerns about the accessibility of a Virtual Reality (VR) system she worked on: *"A really famous VR software at the time, had done inclusivity in terms of the color of the skins of hands, right, and allowing for people with one hand to operate it. [...] I brought it up as an option, because we were working with folks who could be helped by that."* (I11), but this was not pursued, and she was told *"well, nobody asked for it."* (I11)

4.1.3 An entire product. Practitioners also surfaced ethical concerns about entire products (e.g. a piece of user-facing software at a SaaS company, a contract with a specific client at a consulting firm, etc.). Unlike feature-level concerns, product-level harms could not be remediated through minor implementation changes. When the product's very existence is the issue, concerns can only be resolved when the product is shut down or its underlying business logic radically altered.

For instance, one practitioner working in marketing consulting firm was assigned to develop a customer segmentation model for selling high interest loans, a practice they believed would harm borrowers: *"developing a machine learning model to predict where to find customers that were likely to both take on unsustainable amounts of debt and to be able to pay it back" (I5)* In this case, no amount of tweaking the customer segmentation model would have chipped away at the irreconcilable issue the practitioner took with the project — their core belief that high interest loans are inherently harmful to customers.

Another interviewee reported being assigned onto a project to make improvements on telecom software which he was concerned were 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. And kind of 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.

4.1.4 An organization's raison d'être. Finally, some practitioners reported irreconcilable concerns with their organization's goals, core products or business practices. Practitioners often became disillusioned with their organizations when they developed a belief that their organization or industry's core mission is harmful.

Many practitioners brought up concerns that their work was being put to military ends, constituting the most common concern type (see Table 2). These included concerns of direct harm *"the software I was contributing to was being used to harm innocent civilians or infringe on human rights" (S174)*, but also ideological issues, such as *"am I indirectly contributing to the ills of imperialism?" (S161)*. One practitioner cited his newly-held Buddhist faith as the origin of his concerns, saying *"working in the weapons domain is actually really not good karmically" (I1)*, but later reflected that *"if you 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, who was hired onto a Fintech firm to combat fraud, had issue with his day-to-day but instead with his firm's overall business: *"Preventing [fraudulent users] 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 that the core service offered by the organization was harmful, therefore concerned about the very reason the company existed. He reflected on how being so fundamentally misaligned with his company over business ethics 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)*

Practitioners also voiced organization-level concerns about consulting firms. One former consultant we interviewed voiced concern with his previous firm's lack ethical standards in chosing which clients to work with: *"The company, as it turns out, also does a fair amount of work for [...] oil companies, [...] firearms, [...] British American Tobacco [...] not exactly paragons of morality." (I5)*.

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

In response to being asked to do work they considered unethical, the vast majority of respondents (87 of 115) took actions to resolve their concerns, though most felt their concerns were not unresolved (62 of 87). Some practitioners saw their actions as operating within the proper escalation channels of their organizations, while others adopted "tactics of resistance" [91]: working within their powerful institutions in ways that unsettled or diverged from its prescribed conventions [19]. Still others contextualized their actions outside their workplace altogether, focusing on their ability

to affect change as skilled tech workers, organizers, or politically engaged citizens. In this section, we will delineate the various actions practitioners, discuss when certain actions are effective and present challenges practitioner reported with acting on ethical concerns.

5.1 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 ethical concerns (see Figure 1), because harms resulting from the core purpose of a product 4.1.3 or the business practices of an organization (*i.e.*, those in sec. 4.1.4) are not able to be resolved through changes to system implementation. Furthermore, even when practitioners see opportunities for technical solutions, their actual implementation is contingent on whether management agreed that harms were important, and willingness to expend resources to fix them.

Both interviewees I6 and I10 (whose concerns were summarized in 4.1.2) came up with technical solutions that would have resolved their concerns, which were denied by management for different reasons. 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 [...] and then turn the tracking on if they're in. If they're not, the tracking is off." (I6)*. He reported that management was sympathetic towards his concern over privacy: *"[my manager] agreed with the ethical standards that we have [...] we cannot monitor people's comings and goings" (I6)*. However, 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 human in the loop solution to his concern of denying people insurance coverage due to GPS rounding errors: *"we [could] have a three value response instead of a binary yes no response [the third being] "maybe need to check further if it was right on the boundary"" (I10)*. However, what the practitioner saw as a serious harm *"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. This disagreement over what constitutes "harm" led the client to dismiss the concern entirely: *"the client cut me off and told me she didn't care and that that she just that I just needed to do it." (I10)*. The latter practitioner 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)*

Proposing a solution to rectify a perceived harm was often already outside of the assigned responsibilities of those in independent contributor (*i.e.*, non-management) roles, especially for junior practitioners: *"I think whatever was given to me [...] by my product manager was what was in my control [...] that could be biased because I am a junior candidate." (I15)* This interviewee described feeling agency only on matters of implementation within already scoped assignments, so suggestions to rework aspects of a product in response to ethical concerns is a substantial renegotiation of his responsibilities.

5.2 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.*, which individual is in this picture?) reported successfully pivoting the direction of his project to facial verification (*i.e.*, are these pictures of the same individual?). He raised ethical concerns about downstream harms like bias and surveillance to management, but also couched these within organizational incentives to pursue easier and more achievable projects (verification)

before difficult and experimental ones (identification): *"because we were understaffed [...] was also the motivation for me to say, 'Okay, let's not focus on something that is ethically concerning [...] we don't have the resources to do it.' [...] I raised the ethically concerning problems, but there was also an argument for not going for it because we didn't have enough people to work on it." (I14)* He voiced his ethical concerns, but also used institutional incentives by justifying his desired outcome as easier to implement. 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" [57] — 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 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)*

5.3 Refusing Assignments

One common action respondents self-reported was refusing to work on the task they found unethical. Refusals varied — some threatened to quit loudly, while others resigned quietly; some sought reassignment at their current firms, while others went to great lengths to move to another industry.

5.3.1 "Quiet Quitting". Some practitioners did not explicitly refuse their assignments, instead quietly reducing their output to a minimum. 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 5.4). 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)*.

5.3.2 Reassignment. A handful of respondents reported staying at their organizations even after refusing to work on projects they were concerned about. These practitioners were able to remove 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 also typically needed some seniority to ask for a reassignment, as one participant describes: *"My seniority and wide swathe of other projects to choose from." (S19)*

Company policy made it easy for another respondent concerned about his work with a weapons manufacturer to be reassigned: *"[I] talked to my supervisor. We had a policy at the company that nobody has to take part in any software projects involving military use." (S95)* In contrast, an interviewee at a different firm had a completely different experience: *"They were insisting that I [...] switch over to the defense contract. And I told them, I didn't want to do that. [...] I don't know if [...] they needed to bill hours to the defense contract, or if they really just needed the help." (I1).*

5.3.3 Ultimatum of Quitting. A handful of practitioners described delivering ultimatums to management – putting their job on the line and making it clear that they would quit unless some harm was rectified, 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 earn the hospital a bonus but jeopardize patient outcomes. In response, they raised the concern to management and put their job on the line: *"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.

Sometimes, an ultimatum coming from a practitioner is a wake up call for management, forcing them to take seriously harms they may have neglected in the past, as one participant said: *"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 at all. I think [...] there are a lot of engineers that [haven't] demonstrated they're worth a lot to the company. So it kind of depends on the individual, whether you have leverage over leadership." (I7).*

5.3.4 Actually Quitting. : 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 asked to work on spam software 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, we found cases where if a crucial developer leaves an already precarious project, this can cripple and stop 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)*

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

The strategy of "feet voting" describe 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 and tried to find a different job" (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. And so I thought, well, like, 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: *"After I made the mistake, the second time, 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: *"[others] absolve themselves of responsibility, because [...] their livelihood is threatened. But [...] that's really hard argument [for software developers] to make. [...] 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 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, more often, participants did not feel this way (see 6.1).

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 [46]. 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)*

5.5 Leveraging legal systems

One practitioner we spoke to attempted to collect information to raise his concern with law enforcement: *"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 6.3), 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.*, [88]) license in order to limit downstream harms of OS agricultural software he created, but conceded that it was unlikely that he would have the resources for costly litigation to enforce them. In discussing the efficacy of his action, he compared the process of choosing a license to what he 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)

5.6 The psychological toll of raising concerns

Acting on one's concerns is easier said than done. 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 practitioners employed full-time, for whom their organizations are often 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 escalating ethical concerns to a client as: *"one of the most terrifying moments in my life."* (I10)

Practitioners 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 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, really, because [...] by me saying, I'm not willing to do that, you know, I know other people take it as a moral judgment against them."* (I1) A survey respondent echoed this feeling of judgment, suggesting that discussing workplace ethics with colleagues could backfire and damage trust: *"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) On the other hand, practitioners described socializing and circulating concerns among peers in an attempt to feel less isolated in grappling with concerns. For instance, one interviewee leveraged their organization's employee directory and intranet to *"find other people who cared about the same things."* (I4) They started attending reading groups, and eventually transitioned into a more ethics focused role. Looking back, they reflected: *"Finding community in the ethical AI space made me feel so much more grounded in understanding these issues."* (S14). One interviewee makes the case that these "watercooler" conversations help break the patterns of silence and provide foundations for future action: *"Most of the time, people would just say nothing, and it's moot, right? [...] just saying that idea, I think is a good first step."* (I15)

The aftermath of a failed escalation can also seriously affect practitioners' mental health. One interviewee recalled his state after a failed escalation attempt: *"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 [and I] had to actually force myself to communicate with work."* (I5) Another interviewee described similar emotions after his concerns were dismissed by both the client and his direct manager: *"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)

Disagreements over ethical concerns are psychologically taxing, since they challenge practitioners' beliefs about right and wrong. The possibility of jeopardizing their immediate livelihood and longer career makes the process doubly distressing. All the while, the process of raising concerns erode practitioners' support structures, making it even harder to take action for uncertain impact.

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

In the sections above, we presented software practitioners' ethical concerns, how certain actions map onto concerns of different scopes, and the challenges of each action. In this section, we discuss personal and organizational factors which affect practitioners' ability to see their concerns satisfactorily resolved. We found that financial and immigration precarity, company culture, and organizational incentives were major factors influencing practitioners' ability to resolve their concerns.

6.1 Financial and Immigration Precarity

While some software engineers felt comfortable turning down jobs (Sec. 5.4) or quitting their current jobs (Sec. 5.3.4) 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? Everybody has their personal Maslow hierarchy of needs [...] 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) (participant referencing [53]). 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 [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 [27]) 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, saying his 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 practitioner faced this 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)

6.2 Workplace Culture

Organizations with similar core products or business models may have radically different workplace cultures that affect practitioners' feeling of agency, and even resolution outcomes. We use the term *culture* to encompass an organization's norms, expected practices, and communication styles.

Respondents cited transparent and collaborative cultures as factors that made them feel more comfortable acting on their concerns at organizations. In response to a question asked if anything made taking action feel easier, respondents wrote: *"A general environment of trust and respect, a common goal of making a quality product." (S71)*, *"'Open door' policy and small company size making it easy to get 1-on-1 time with execs." (S52)*. However, other respondents describe *"hostile"* *(S72)*, *"authoritarian/passive aggressive management style[...] hierarchical culture" (S110)*, *"suggestions from higher-ups that ethics discussions were a waste of time." (S52)* and as things that made it more difficult to act.

Interviews illustrated practitioners' views on the link between culture and ethics. For example, one said: *"There's a culture of openness and transparency here [...] from the CEO level downwards [...] we have acknowledgment and that that's a value." (I7)* Another practitioner (I10) contrasted two consulting experiences, one in which he was isolated from his colleagues in cubicles, where his concerns were unresolved and he resigned, and one where he worked closely with his colleagues and his concerns were successfully resolved after he felt able to bring up his concern.

A nuance about culture we observed is that practitioners did not necessarily feel more comfortable voicing concerns in organizations with "friendly" cultures. In fact, several survey respondents mentioned that having friendly, personal relationships with management made it harder to act on their concerns, in responses to the questions about things that made it feel harder to act on concerns: *"I knew the founder quite well" (S94)*, *"The boss was a friendly chap" (S104)*, *"The social interaction and bonding attempts from the owners" (S68)*, *"The person who recommended me for the position was my friend and my boss." (S43)* Practitioners who are socially immersed in their workplaces may feel that raising ethical concerns comes across as a moral indictment against their colleagues. Along these lines, one interviewee said that working remotely and having less social ties made it easier to escalate his concerns: *"This remote way of working [...] helped me to create this barrier [...] this disconnection with the manager helped me to say [...] I care less about your opinion on this. So it helped me to think more rationally." (I14)*

6.3 Organizational Incentives

Participants demonstrated an acute awareness of their organization's incentives, and used them to reason about their power to act on their ethical concerns: often the root cause of ethical concerns, though rarely the framework through which concerns were resolved.

Profit motives often both framed ethical discourse, and lead to ethical concerns. Many survey respondents made this link explicitly, for example: *"[...] features were implemented to earn money by any means necessary" (P69)*, and: *"they were selling geolocation data because it's worth a lot of money." (P67)*. One interviewee suggested that firm's financial struggles leads to ethical concerns: *"in many businesses [I've seen that] between the choice of closing the business, and a choice of doing something uncomfortable, almost everyone chooses to do something uncomfortable" (I7)*, recalling when his previous employer sold user data to advertisers to survive a revenue downturn, saying *"[we were] scrambling like, is there any way [to] make more money [...] some new revenue stream?" (I7)* One practitioner described how he was pushed to work on a military contract so his company could bill as many hours as possible: *"the way government contracts work is they'll just pay for heads to be on a project [...] The companies see it, like that's easy money" (I1)* As seen in 5.2, initiatives to mitigate harms often had to be couched within incentives like revenue or user growth to gain support. One interviewee described his organization using "ethics wins" for brand image gains: *"It's become a thing that you brand yourself as a diverse organization [...] it's become not an ethical concern, but a marketing concern, to be honest. And the way that incentives align." (I2)*, later suggesting that companies sometimes accept ethical commitments that hurt their bottom line in the short term to boost revenue in the long run.

Other practitioners suggested that government agencies, non-profits, and companies can have values that transcend profit. One interviewee who recently transitioned into public service reflected on the difference as *"you're pursuing your goods beyond profit, right? [...] versus 'we want to make money'" (I5).* Another computer science researcher 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. So we're not looking for what's good for the enterprises. But what's good for the people?" (I20).* However, other academics described how pressures to publish lead to concerns over methodologically and research integrity (see Table 2).

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)*

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

Our work shows that even highly paid, mobile, in-demand and ultimately *powerful* software engineers [17] – have less power than one might think, as power is still a limiting factor in their ability to resolve their ethical concerns. Identifying ethics concerns is only half of the struggle, and an unfulfilling one without the ability to ensure they are resolved. Ethics interventions such as toolkits [48, 70], checklists [52], principles [42], and education [22] are often designed to help practitioners identify issues, and flag them to others, for example using model cards [59] or datasheets [26]. However, these interventions practically depend on practitioners having power to dedicate resources, make design changes, cancel concerning features or products, or otherwise fix issues these interventions may help identify. Without this power, these interventions risk being insufficient at best. At worst, they risk limiting critique to the narrow scope of system design and allowing companies to avoid substantive scrutiny of business practices [31], enforceable regulations [65, 81], or fitting into a simple narrative where a moral failing of tech workers is the core problem and better ethics education the solution [95]. Our work shows what happens after practitioners identify concerns without the help of (and often, exceeding the scope of) these ethics interventions – and discover severe limits on their power to affect change as they attempt to resolve their concerns. In this section, we recap how these limits depend on the scope of a practitioner's concern, and discuss similarly dependent implications for future tech ethics research and education.

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

On the left side of our spectrum (Figure 1) are "bugs": where a technical fix is possible, at least in theory. Ethics toolkits and checklist are not set up to raise or resolve concerns beyond a narrow scope of ethics issues on the left side of the spectrum of issue scopes shown in Figure 1, and even then, they depend on practitioners having the power to ensure resources are dedicated to fix any identified issues. Our work shows that they may be ineffectual if they do not account for the overriding role that power plays in engineers' ability to see their concerns resolved. Interventions such as Fairness checklists [52], model cards [59], and dataset datasheets [26] may "empower[...] individual advocates" to talk about ethics [52], similarly to how our interviewee cited papers on inclusivity in VR in section 4.1.2, but she

was nonetheless unsuccessful in persuading her team to dedicate resources to these concerns, given that clients had not demanded it. The incentives won out, even for a concern that did *not* challenge the *raison d'être* of her product or company, and that indeed was aimed at improving the design of her product rather than critiquing its entirety..

In these cases, interventions such as checklists may empower advocates to raise fairness issues [52], and personas may help identify and raise "uncomfortable design discussions" [37] about gender bias in software design [14, 14]. These tools legitimize ethics concerns, in part by framing them using the more palatable critique of improvements to a product (*i.e.*, as "bugs") to improve its chances of success [40, 91]. In this way, some of our participants tactically raised their ethics concerns in terms of their organization or team's incentives (see 5.2) and are sometimes successful (*e.g.*, crane software in Sec 4.1.1). However, we show that even when these less-threatening, narrowly-scoped issues may garner agreement that a concern is legitimate (a "bug"), these concerns are often nullified using the usual logics of "customer centricity" (*i.e.*, something the customer did not demand, *e.g.*, inclusive VR in 4.1.2) or that the client or management is unwilling to dedicate resources to remedy (*e.g.*, two examples in Sec. 5.1). In a well-known outside example, Google researchers raised concerns about, and outlined avenues to remedy, bias and the carbon impact of large language models [8], yet Google received this particularly poorly [56].

In view of the centrality of power, our work helps provide "guidance around how to navigate organizational power dynamics" [92] when proposing solutions to ethical concerns that toolkits may identify, by helping understand the power structures into which ethics interventions may fit, and the limits on the power of those who may apply them. We find that fixing "ethics bugs" often relies on practitioners' *power* to persuade others to dedicate resources to fixing them (see Sec. 5.1), and this motivates further work to develop tactics of persuasion such as justifying solutions to ethics problems in terms of organizational incentives (see Sec. 5.2, see also [91]), and work to quantify and provide outside evidence for relationship between ethics fixes (*i.e.*, accessibility, see Sec 4.1.2) and the incentives that decision maker care about, such as product success or user growth. Additionally, as some empirical software engineering research shows [32], practitioners may themselves have power to prioritize among bugs and our work, suggests the opportunity to examine where "ethics bugs" lie in their prioritization.

Other work suggests that concerns over career progression limit practitioner's power to advocate for ethics fixes [52, 69], our work shows what happens when practitioners accept this risk to the point of being blacklisted, but also how this power is limited by financial and immigration precarity (see Sec. 6.1, see also [12]) or enabled by workplace culture (see Sec. 6.2), suggesting future research into other contingencies on practitioner's power to advocate for ethics, and in other contexts. Additionally, for fear of a checklist or toolkit foreclosing on the kind of self-identified concerns we study that may not be enumerated, design research could focus on ways to ensure these interventions are received not as exhaustive inventories but instead as starting places for open-ended critique (see [16] for discussions of relative merits of open *vs* closed mechanisms of critique in another context).

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" [22], but fewer than one quarter touch on the systems of power – "capitalism, financial models, marketing, pricing" – within which issues must be addressed. Our work suggests that tech ethics courses ought to broaden their emphasis to include teaching students an awareness of these and other power structures which await them in their future careers (and indeed, presently surround them in their education and internships), and tactical skills to raise their ethical concern, lest courses help identify issues but leave them unprepared to advocate for fixes. Indeed, even awareness of these power structures may in turn engender awareness of new issues – as privilege can insulate from consideration of downstream harms [88]. Our work also

enumerates ethical concerns that practicing software engineers face (see Table 2), which could help ensure in class examples prepare students for the ethical concerns that practicing software engineers face at work.

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

The right side of the spectrum in Figure 1 characterizes practitioner concerns which question the *raison d'être* of their organization or industry, for example when a practitioner working at a military contractor develops concerns about the way the military uses their technology. Others have critiqued design-stage interventions as insufficient [25], especially when harm is inherent in how systems are used [88], 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" [64]. 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. 5.3), and 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 familes [2, 9], but the contract continued. However, we hear from some participants that resigning is disruptive, and even leads clients to cancel precarious projects (see Sec. 5.3). Future research can explore individuals' power over outcomes: when is a resignation by a concerned engineer successful in halting a software project? Research in volunteer open source communities identifies "truck factor" developers whose departure puts the project into serious peril [4], future work can explore this in for paid developers in company contexts. Some of our participants practice "feet voting" by proactively planning their career in alignment with their (see 5.4), future work can evaluate how commonly and with what priority ethical concerns factor into tech job seekers' priorities, and examine support or information needs for ethically-concerned job seekers.

How might tech labor organizing account for these larger critiques? If software engineers working in, for example, military technology, believe their industry is fundamentally immoral, are there opportunities for collective action to shape or reform (of any variety, reformist or non-reformist [30]) how their technology is used, before outright refusal? Tech workers unions, while advocating for ethics issues in addition to more traditional union issues like working conditions [1, 17], may nonetheless be reluctant to embrace wholesale critiques of their industry if doing so becomes so successful it leads to layoffs.

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" [22]. Courses may consider such exercises, along with linking this to their career planning a primary opportunity to align their labor with their values (see 5.4). Tech ethics education may also consider teaching about tech worker rights organizations such as the Tech Workers Coalition [1]. Tech education can also to expand encompass skills we report our participants using, such as negotiating for ethics using organizational incentives (see 5.2), and the more strategic power of "voting with one's feet" to make ethics a factor in how one plans their career (see 5.4). It can also call attention to strategies for building collective power, including watercooler talk to socialize concerns (see 5.6), whistle blowing and legal remedies (see 5.5) to discussion of tech worker unions (see 5.4).

7.3 Who is organized? The coherence of focusing on "AI" or "Big Tech" in tech ethics discourse

As one of our participants recognized, "the work doesn't get done without us" (see 5.4), and the power of labor need not limited to individuals, but can be collective. However, while some of our participants attempted "water-cooler talk" to attempt to find support for their concerns (see Sec. 5.6), only two of our participants raised unions as an avenue for collective action on ethics concerns (see Sec. 5.4) despite despite high-profile efforts to collectively organize over ethics issues at large firms such as Google and Microsoft [46]. We argue that our results show that tech ethics research ought to: firstly, broaden to consider ethics generally instead of a focus on AI; and secondly, ethics research oriented at understanding practitioner needs ought to broaden away from software engineers at "Big Tech" companies. This larger focus will examine more contingencies in tech worker power, help build a broader coalition by recognizing issues of overlapping concern, and also help shift consideration of ethics towards harm irrespective of implementation instead of a privileged focus on "AI" harm.

7.3.1 Ethics beyond AI. Firstly, to capture as wide of a scope of ethical concerns as possible, and given divergent conceptions of what "AI" is [45], we did not limit our study to "AI" practitioners, or to concerns related to "Ethical AI". While one analysis found that "activism" by "the artificial intelligence (AI) community" had been largely successful in part because of "a coherent shared culture" where people "publish at and attend the same major conferences", and concluded that "The AI community is acting together – it is organised" [7], we argue that considering activism from the "AI community" without explicitly defining it explicitly risks casting it as a monolith, characterized by its most privileged members, and the sources of power and concerns they hold, and their unique context. Despite not deliberately recruiting AI practitioners, most of our interviewees were "building 'smart' machines" (*i.e.*, as used in [18]), and some explicitly positioned themselves as working on "AI" systems. Despite this, none of our "AI" participants consider themselves "organized" (though one had attempted organizing, see Sec 5.4), or referenced any of these conferences, and few talked about themselves as part of a wider shared culture indicating less coherence than these statements in [7] suggest.

We argue that our results show that considering "AI" when considering ethics concerns adopts a limited scope of scrutiny, thereby again, focusing on design-stage interventions [25]. Using "AI" is a design choice, whereas many of the concerns our participants raise do not depend on whether the specific system used for harm uses "AI" or not (especially true for concerns with the *raison d'être* of their industry, see Sec 4.1.4). We therefore argue that future work on the ethics challenges software practitioners face 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 [31]. We also suggest that our scopes in Figure 1 may be a way to conceptualize different kind of practitioner concerns that is not limited to AI. Similarly, in an educational context, pedagogy on "AI Ethics" [10] may consider broadening to study tech ethics more generally, as many courses already do [22].

7.3.2 Ethics beyond Big Tech. Secondly, only one of our interviewees currently works for a "Big Tech" company (*i.e.*, [72]) though he did not speak of concerns working there, and only one spoke of their 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 Table 1). This is relevant in light of calls to do research beyond "large, internal software development teams" [77], but also by claims that certain AI 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." [17]. Major companies invest heavily in AI resources and (certain farmings of) ethics to the point they raise concerns of

capture of not only AI resources [85] but also AI ethics discourse [94]. These companies also are the sites of the most high-profile examples of unions and other collective organizing [3, 15, 46], and thus their workers may well 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" [91] or "major companies" [69] risk assuming a base level of exposure to (certain framings of) AI ethics discourse, and thus risk assuming a certain level of generality around what ethics concerns exist. Our work suggests that AI Ethics research may need to broaden to better account for the majority of software practitioners who do not work the small number of "major" companies. For example, we believe that our participants' feelings of being isolated in their ethics concerns and resulting mental health consequences (see Sec. 5.6) and attempts to build this community by socializing their concerns (see Sec. 5.6) may reflect unique isolation in contrast to in tech-centric companies, where processing of ethics concerns with similarly aware colleagues may help [76]. 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. For example, this could proceed using one database which catalogs 36 tech labor unions and 506 examples of collective actions in tech [63] including those by warehouse blue collar workers, white collar software engineers, and actions spanning both, in concert with calls for HCI research to support all kinds of tech labor [23].

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] Oliver Bates, Kathy New, Samantha Mitchell-Finnigan, Matthew Louis Mauriello, Christian Remy, Roy Bendor, Samuel Mann, Simran Chopra, Adrian K Clear, and Chris Preist. 2019. Towards a responsible innovation agenda for HCI. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–8.

[6] Andrew Begel and Thomas Zimmermann. 2014. Analyze this! 145 questions for data scientists in software engineering. , 12–23 pages.

[7] 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

[8] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In *2021 ACM Conference on Fairness, Accountability, and Transparency*. 610–623.

[9] 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/.

[10] 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

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

[12] 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

[13] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. PMLR, 77–91.

[14] 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

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

[16] Lars Thøger Christensen, Mette Morsing, and Ole Thyssen. 2017. License to Critique: A Communication Perspective on Sustainability Standards. *Business Ethics Quarterly* 27, 2 (April 2017), 239–262. https://doi.org/10.1017/beq.2016.66

[17] 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

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

[19] Michel De Certeau. 1984. The Practice of Everyday Life. 1984. *Trans. Steven Rendall. Berkeley: U of California P* (1984).

[20] 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

[21] 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).

[22] 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

[23] Sarah E. Fox, Vera Khovanskaya, Clara Crivellaro, Niloufar Salehi, Lynn Dombrowski, Chinmay Kulkarni, Lilly Irani, and Jodi Forlizzi. 2020. Worker-Centered Design: Expanding HCI Methods for Supporting Labor. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–8. https://doi.org/10.1145/3334480.3375157

[24] Batya Friedman. 1996. Value-sensitive design. *interactions* 3, 6 (1996), 16–23.

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

[26] 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.

[27] 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.

[28] 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).

[29] 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.

[30] Ben Green. 2021. Data Science as Political Action: Grounding Data Science in a Politics of Justice. (2021), 17.

[31] 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*.

[32] 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.

[33] Hana Habib, Megan Li, Ellie Young, and Lorrie Cranor. 2022. "Okay, whatever": An Evaluation of Cookie Consent Interfaces. In *CHI Conference on Human Factors in Computing Systems*. 1–27.

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

[35] 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

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

[37] 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

[38] Michael Hilton, Nicholas Nelson, Timothy Tunnell, Darko Marinov, and Danny Dig. 2017. Trade-offs in continuous integration: assurance, security, and flexibility. , 197–207 pages.

[39] 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.

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

[41] 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

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

[43] 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.

[44] Lorraine Kisselburgh, Michel Beaudouin-Lafon, Lorrie Cranor, Jonathan Lazar, and Vicki L. Hanson. 2020. HCI Ethics, Privacy, Accessibility, and the Environment: A Town Hall Forum on Global Policy Issues. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3334480.3381067

[45] 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

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

[47] 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.

[48] 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.

[49] Yvonna S Lincoln, Susan A Lynham, and Egon G Guba. 2011. Paradigmatic controversies, contradictions, and emerging confluences, revisited. *The Sage handbook of qualitative research* 4 (2011), 97–128.

[50] 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.

[51] 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.

[52] 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.

[53] A Malsow. 1943. Hierarchy of needs: A theory of human motivation. *Psychological Review, 50, 370-396* (1943).

[54] 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.

[55] 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.

[56] Cade Metz and Daisuke Wakabayashi. 2020. Google Researcher Says She Was Fired Over Paper Highlighting Bias in A.I. *The New York Times* (Dec. 2020).

[57] Debra E Meyerson. 2008. *Rocking the boat: How tempered radicals effect change without making trouble.* Harvard Business Review Press.

[58] Craig S Miller. 2011. Item sampling for information architecture. In *SIGCHI Conference on Human Factors in Computing Systems*. 2211–2214.

[59] 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.

[60] Brent Mittelstadt. 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence* 1, 11 (2019), 501–507.

[61] Dawn Nafus and Jamie Sherman. 2014. Big data, big questions| this one does not go up to 11: the quantified self movement as an alternative big data practice. *International journal of communication* 8 (2014), 11.

[62] Anton J Nederhof. 1985. Methods of coping with social desirability bias: A review. *European journal of social psychology* (1985).

[63] Nataliya Nedzhvetskaya, JS Tan, Hyatt Dirbas, and Wynnie Chan. 2022. Collective Action in Tech. https://data.collectiveaction.tech/.

[64] 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

[65] Rodrigo Ochigame. 2019. How Big Tech Manipulates Academia to Avoid Regulation. https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/.

[66] Will Orr and Jenny L Davis. 2020. Attributions of ethical responsibility by Artificial Intelligence practitioners. *Information, Communication & Society* 23, 5 (2020), 719–735.

[67] 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.

[68] 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

[69] 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).

[70] 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

[71] Marc J Riemer. 2007. Communication skills for the 21st century engineer. *Global J. of Engng. Educ* 11, 1 (2007), 89–100.

[72] 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

[73] 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.

[74] Reginald G Smart. 1966. Subject selection bias in psychological research. *Canadian Psychologist/Psychologie canadienne* (1966).

[75] Anselm Strauss and Juliet M Corbin. 1997. *Grounded theory in practice.* Sage.

[76] 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.

[77] 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.

[78] 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.

[79] Jeroen Van den Hoven. 2013. Value sensitive design and responsible innovation. *Responsible innovation: Managing the responsible emergence of science and innovation in society* (2013), 75–83.

[80] 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.

[81] Ben Wagner. 2018. Ethics as an escape from regulation. From "ethics-washing" to ethics-shopping? (2018).

[82] Jennifer Wang and Angela Moulden. 2021. AI Trust score: A user-centered approach to building, designing, and measuring the success of intelligent workplace features. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–7.

[83] Robert Philip Weber. 1990. *Basic content analysis*. Number 49. Sage.

[84] Robert S Weiss. 1995. *Learning from strangers: The art and method of qualitative interview studies*. Simon and Schuster.

[85] Meredith Whittaker. 2021. The steep cost of capture. *Interactions* 28, 6 (2021), 50–55.

[86] David Gray Widder, Laura Dabbish, James D Herbsleb, Alexandra Holloway, and Scott Davidoff. 2021. Trust in Collaborative Automation in High Stakes Software Engineering Work: A Case Study at NASA. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13.

[87] David Gray Widder, Michael Hilton, Christian Kästner, and Bogdan Vasilescu. 2019. A conceptual replication of continuous integration pain points in the context of Travis CI. , 647–658 pages.

[88] 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*.

[89] Kelly Widdicks, Bran Knowles, Gordon Blair, Carolyn Ten Holter, Marina Jirotka, Federica Lucivero, Gabrielle Samuel, and Helena Webb. 2021. Anticipatory Governance in the Technology Sector: Processes, Critiques and Principles for Addressing Grand Challenges in Computing. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–5. https://doi.org/10.1145/3411763.3441314

[90] Langdon Winner. 1980. Do artifacts have politics? *Daedalus* (1980), 121–136.

[91] 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.

[92] Richmond Y. Wong, Michael A. Madaio, and Nick Merrill. 2022. Seeing Like a Toolkit: How Toolkits Envision the Work of AI Ethics. arXiv:2202.08792 [cs]

[93] 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).

[94] 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

[95] 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 [84], with 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 as a [ROLE] 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?