

Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics

Jacob Metcalf (Data & Society), **Emanuel Moss** (Data & Society, CUNY Graduate Center),
danah boyd (Microsoft Research, Data & Society)

INTRODUCTION

Ethics is arguably the hottest product in Silicon Valley's hype cycle today, even as headlines decrying a lack of ethics in technology companies accumulate. After years of largely fruitless outside pressure to consider the consequences of digital technology products, the very recent past has seen a spike in the assignment of corporate resources in Silicon Valley to ethics, including hiring staff for roles we identify here as "ethics owners." In corporate parlance, "owning" a portfolio or project means holding responsibility for it, often across multiple divisions or hierarchies within the organization. Typically, the "owner" of a project does not bear sole responsibility for it, but rather oversees integration of that project across the organization.

A remarkable range of internal and external challenges and responses tends to fall under a single analytic framework called "ethics." This strains an already broad term that in some contexts means an open-ended philosophical investigation into moral conditions of human experience and, in other contexts, means the bureaucratized expectations of professional behavior. Likewise, it places strain on corporate structures because it is bureaucratically challenging to disambiguate whether these problems belong in the domain of legal review, human resources, engineering practices, and/or business models and strategy.

One of our informants for the project described in this essay illustrates what it looks like to "own ethics" in the tech industry today. She works for a global enterprise-software provider—a company that serves software platforms to other companies—headquartered in Silicon Valley, and her job title refers to ethics. Charismatic and gregarious, she is comfortable talking to engineering staff, corporate heads, and external industry critics, and is known to give hugs to new and old acquaintances, as is common in Northern California. While always friendly, she is often blunt with her judgment, and despite decades of engineering and management work, she is fluent in the discourses of justice, equity, and fairness. She ably uses those discourses to explain to corporate leaders and the public where her industry has gone wrong. On any given day, one might find her giving a TED-style talk about ethical design practices at an industry conference, running a closed-door workshop for fellow ethics owners and haranguing attendees to write ideas on sticky notes for the whiteboard, writing an academic conference paper about tech ethics, or conducting internal product oversight. Internally, she is known as a prolific evangelist for ethics in company communications channels. She often presents her central goal as creating an ethics race to the top across the

tech industry, demonstrating her conviction that tech can do better to fulfill foundational ethics commitments.

All this activity happens against a backdrop of the usual Silicon Valley cultural oddities that contrast with the seriousness of the task: offices with swing sets and beer kegs, conferences run by vaporware venture capitalists, and engineers who attempt to troll our informant on Twitter. In our interview, she repeatedly gestured to her role as someone who translates external norms and pressures into practices that are internally tractable—for example, rendering the UN’s Universal Declaration of Human Rights as a practical guideline for screening out problematic enterprise clientele, or finding ways to align revenue-generating metrics (“clicks”) with an ethically robust model of value for platform users. Upon reflection, it is remarkable that a profit-seeking enterprise would employ someone in this role who spends a fair amount of her salaried time criticizing the industry’s normative failings.

This informant is one amongst the growing cadre of people tasked with owning ethics who lead us to question whether and how ethics, as an organizational responsibility, can challenge the core logics of the industry that repeatedly animate its ethical crises. Through ethics owners, the ancient, domain-jumping, and irresolvable debates about human values that underlie ethical inquiry are implemented as institutional practices aimed at technology development. While a robust effort to foreground the importance of human values is both welcome and needed, a closer examination of how ethics is becoming institutionalized as a set of roles and responsibilities, and operationalized as a set of practices and procedures, reveals the risk of a premature foreclosure of the fundamentally open-ended and irresolvable questions that underlie human value commitments.

Given the increasing power and centrality of artificial intelligence (AI) and automated decision-making tools in everyday life, there is an urgent need for a coherent approach to addressing ethics, values, and moral consequences. This need is complicated by the centrality of personal data to many of the tech industry’s products, its own discourse presenting itself as “making the world a better place,” and the lack of a timely regulatory response to many of the industry’s problems. Ethics owners operate inside a fraught dynamic: on the one hand attempting to resolve critical external normative claims about the core logics of the tech industry; on the other hand, doing so while fully embedded within those logics. Attempts to institutionalize ethics within entities structured by core logics of corporate power point towards a series of structural, conceptual, and procedural pitfalls that may ultimately stymie these efforts.

METHODOLOGY

This paper draws on ethnographic, textual, and other qualitative data that the authors have collected as part of a broad engagement with the development of data-driven technologies. These data were supplemented with interviews with 17 people from established and well-known companies who either “own ethics” in their formal role or who have made addressing ethics within their companies or the industry a personal mission.

In our interview corpus, two individuals are outspoken tech workers without supervisory responsibilities; the others are team leaders or principals who have some aspect of ethics ownership within the ambit of their job title. All of them work (or worked) at companies that have more than 1,000 employees and are headquartered or have sizeable offices in California. Most have strong technical backgrounds in computer science, although a few also hold MBAs or social science degrees, and one has a humanities background. These interviews complement a range of ethnographic and interview work that the authors independently conducted prior to and during this project.

SITUATING ETHICS IN THE TECH INDUSTRY

The current trend of Silicon Valley corporations deciding to empower ethics owners can be traced to a series of crises that have embroiled the industry in recent years. These crises appear to the public as spectacular revelations about real or potential harms that technology has produced, as seen in the role of Cambridge Analytica in the 2016 US presidential election and the Brexit referendum; racial bias in the error rates for pre-trial flight risk assessments and facial recognition technology; and the development of tools intended for the battlefield and repressive enforcement of immigration policies. Such revelations enter public awareness through the efforts of journalists and scholars, as well as concerned employees working inside tech companies, but it is not always immediately clear what the ethical stakes of these crises are. Indeed, “ethics” is a capacious term, holding a range of meanings indexed to context of use, and the ambiguity of the term is central to the challenge of capturing what it means to “own ethics” in the technology sector. While others in these companies might “own legal,” “own security,” or “own corporate social responsibility,” ethics owners do not benefit from an existing set of practices and evaluative measures to guide their actions—although how such practices became institutionalized may be instructive here—nor are there clear external regulations or requirements driving their approach.

Ethics owners are not working in a vacuum, however. Business consultants, professional organizations, nonprofit organizations, and scholars all contribute to the discourses that shape how ethics are being approached by Silicon Valley companies. Universities have implemented ethics curricula for computer scientists (Fessler 2018); critical voices have called for codes of ethics (Ananya 2016); and corporate boards have released statements of principles (Greene, Hoffman, and Stark 2019). Concerns about ethical engineering and business practices are co-emergent with both increased coordinated expressions in Silicon Valley of labor power, by engineers insisting on the right to veto their participation in projects they object to, and the rise of #me too and #times up discussions about structural sexism and racism. Major industry actors have even begun citing the potential ethical consequences of algorithmic products as a downside risk in their regulatory financial disclosures (Simonite 2019). For example, Google’s parent company, Alphabet, stated in a recent financial disclosure that AI products “can raise new or exacerbate existing ethical, technological, legal, and other challenges, which may ... adversely affect our revenues and operating results” (Alphabet Inc. 2018).

Given the conceptual and organizational strain that ethics poses for the tech industry, what should the work of ethics owners consist of? What should the *doing* of ethics look like? An obvious place to turn would be applied business ethics research literatures—after all, contemporary tech corporations are not the first companies to face ethical dilemmas. Business ethics literatures typically index normative concerns to the need for coordination between individual behavior and organizational goals. For example, Sims (1991) describes a primary method for institutionalizing ethics within organizations as the mutual understanding and alignment of organizational expectations and individual behavior. Other top priorities Sims outlines focus on the personal psychological needs of employees to receive adequate attention and rewards, identifying employees with their work so they are publicly accountable for ethical lapses, and making public and internal expectations clear.

The alignment of personal psychological states and explicit organizational priorities is a major theme of business ethics. For example, Trevino (1986) shows how employees respond to ethically ambiguous situations in terms of their own stage of cognitive-moral development, and Fleischman et al. (2017) experimentally demonstrate that managers are less likely to intervene in formally unethical behavior of subordinates when they perceive those behaviors as financially benefiting the organization.

These coordination problems are certainly live concerns for ethics owners in the tech industry. For example, the meetings of ethics owners we attended nearly always included discussion of the challenge of getting some metric of “ethics” included in the high-stakes, data-driven annual personnel reviews ubiquitous in Silicon Valley. One ethics owner acknowledged the challenges of developing these incentive structures that are not subordinated to the company’s primary metrics:

[Tech company employees] are incentivized by revenue generated; by launching things; by user engagement like clicks, how long somebody spends on a particular site or app, [all of which are] really about metrics going back to the stockholder—how much value are we adding to our stock and to the stockholder—and values aren’t based on what [we are] doing that’s positive in the world despite whatever values or mottos the individual companies may have.

While *ethics as coordination* is an issue ethics owners grapple with, that framing also misses a crucial element of the current moment: ethics owners are also tasked with “onboarding” critical external perspectives that challenge the core logics through which these businesses have defined success. Nearly all our informants cited this external pressure as a primary reason their roles exist, but they are not working simply to smooth or redirect that pressure, as public relations professionals might. Rather, these crises and the accompanying public pressure have produced an environment that motivates Silicon Valley organizations to devote resources to ethics owners so that they might develop practices that can be deployed within their own organization.

Yet our ethnographic fieldwork, interviews, and textual sources indicate that as ethical product design and governance goals are becoming institutionalized by tech firms, the practices associated with these goals are being crafted and executed according to the existing logics and structures of the technology industry, even as they are responding to outside critiques of these logics and structures. In doing so, they are producing pitfalls that threaten to prematurely foreclose what can be thought or done under the heading of “ethics.” As companies hasten towards practices that can detect crises before they erupt, they also attempt an updated return to “business-as-usual.” This is not to say that business-as-usual cannot be ethical; rather, we point to the potential for the pitfalls introduced by tech logics and organizational structures to prevent the work of ethics owners from addressing the broader goal of a more just algorithmic and data-driven world.

To better understand these pitfalls, we approach ethics and morality as social phenomena and not as primarily philosophical abstractions. The “ordinary ethics” approach has emerged from contemporary debates in ethnographic theory about how to describe the ways ethics and morality structure social life. It attends to how everyday practices reveal the moral commitments embedded in actions, in contrast to the tendency to treat ethics as a form of argument or an abstraction. Lambek (2010) describes ordinary ethics as the search for “the wellsprings of ethical insight deeply embedded ... in the shared criteria we use to make ourselves intelligible to each other.” Ordinary ethics analytically “dissolves” (2010) the ethical into social activity, practice, and judgment, pulling ethical judgment from the transcendental realm of philosophy and into quotidian lived social experience (Laidlaw 2002; Das 2016). It looks to normally unreflective everyday practices and shared linguistic schema to find the normative claims about how the world should be.

Zigon (2007 and 2014) argues that while ordinary ethics ethnography is useful for *locating* ethics in a culture, this method struggles to *recognize* when ethical claims are made as a way to change a culture, as happened in crises like Cambridge Analytica. He emphasizes the importance of the moment of “moral breakdown,” where a mode of life becomes untenable, which opens the possibility of hermeneutic interrogation of everyday moral frameworks. This echo work by Dave (2012) and Mahmood (2012) demonstrating how everyday ethical practices—such as queer activism or forms of religious piety—are leveraged as a form of critique against dominant modes of morality and political power. Establishing a new mode of everyday ethical practice is necessary—because no one can live in a state of breakdown forever—

and ideally would concretize a more just and fulfilling mode of life. As Zigon (2007) puts it, “ethics is the process of once again returning to the unreflective mode of everyday moral dispositions. But this return from the ethical [breakdown] moment is never a return to the same unreflective moral dispositions.”

Dave (2012) similarly defines ethics in terms of a strategic play between resistance to normalization within oppressive moral and political frameworks (“subjections”), and the attempt to creatively foster a more just everyday life: “those practices that emerge from within subjections as a creative, disruptive response to normalization ... [the] narrowing of possibilities to conform to institutionally legitimized norms such as identity, community, national belonging, and the language of law and right” (9).

Despite differences between these theorists, a central lesson of the ordinary ethics model is that *ethics as practice* is foundationally a tension between the everydayness of the present and the possibility of a different, better everydayness (Keane 2017). This dynamic was palpable in interviews, such as when, almost with a sense of exasperation, an engineer remarked,

I don't think there are enough people talking about ethics in the tech industry.... There are people who are impacted by technology who talk about the tradeoffs of technology and how it has affected their lives or their communities.... There are people who have made a career out of being disaffected tech people who regret what they built [even though they] profited from it.

And there are those who are coming from the outside—“professional ethicists or philosophers.” Yet, even though these voices are loud and growing, from her perspective, the vast majority of people in tech “are not yet moved by ethics.”

This engineer was not suggesting that people in tech are immoral or *unethical*. Rather, what she draws attention to is that “ethics” means different things to different people and therefore lacks conceptual and institutional unity. By transposing the ordinary ethics debate to corporate tech culture at a moment like the present, when external critics and, increasingly, many people inside these companies find the current conditions untenable, the everyday practices of these organizations are held up for scrutiny. However, inside companies, ethics owners are tasked with developing strategies to align everyday practices with corporate logics. Our analysis locates the “owning” of ethics in the tension between those methodologically different approaches to describing “ethics.” If the purpose of the ethics owner is to be “moved by ethics,” then ideally they work through this breakdown in order to help return their colleagues to improved everyday conditions.

As outlined by one of our informants who holds a role inside a research department, the job of the ethics owner is determining how to genuinely resolve “sensitive” problems that demand ethical scrutiny while navigating the everydayness of corporate structures, such as managers, cross-functional teams, and review practices, and negotiating the power to green-light a project. One ethics owner from a social media platform company indicated that his job was explicitly to work against the hubristic tendency of Silicon Valley firms to assume that their product teams know how to define and solve problems best; he stated that “when [my team and I] talk to [external] researchers, we spend lots more time listening and asking them questions than we do trying to kind of pitch a core proposition.”

A common theme among our informants was a struggle to identify how their organizations defined ethics and, as a corollary, where technology ethics responsibility should reside within the organizational hierarchy. Nonetheless, we heard many indicate what ethics *is not* and where it *should not* reside. One informant described an internal negotiation about whether a potential client's project was adequately aligned with his organization's goals but defined the negotiation as never really about “the ethical—it was more about the mechanics and politics.” Another defined the ethical as that which is outside the regulatory domains, noting that a lack of regulation made “doing ethics” logistically easier. Informants often cited the challenge of figuring out where in the hierarchy ethics ownership should be placed due to the many existing roles that

seem to be proximate to ethics, such as legal, data for good, corporate social responsibility, safety, content moderation, policy, research, and product design.

There is also often a question of whether the ethics owner should report to an internal-facing executive, such as a chief technical officer or chief information officer, or a more external-facing one, such as a chief executive officer, chief legal counsel, or vice president for product. This indicates that ethics owners are looking to institutionalize and operationalize ethics in the gaps to be found within current structures and practices. In other words, the domain of ethics is being framed as that which the organization is *not yet doing* but is *capable of doing*. However, such gaps do not clearly present ethics owners with a path that is a meaningful departure from “business as usual.” Rather, the available gaps are largely defined by existing structures and practices.

That these projects are working alongside existing logics is demonstrated by the need our informants emphasized for organizational process and senior leadership commitments. Many informants also emphasized that, because of pressure for steep, exponential, “hockey-curve” growth, tech companies are often extraordinarily large before they begin maturing organizationally. By that time, the conversation about ethics isn’t about how to do the right thing but how to avoid downside risk. Financial pressure from investors and the board (comprised primarily of investors) often undermines those who wish to put ethics front and center. Only once a company is confident of its long-term viability might it begin discussing ethics, evidenced by the relative maturity and success of the companies just now investing in formal “ethics” governance. As one of our informants explained, “ethics ... never makes you money but ethics can save you a lot of money.” Similarly, another informant pointed out that as mature organizations begin to take ethics seriously, responsibility moves away from “the data science and engineering managers and directors who are doing it for the right reasons or for the ethical reasons” towards the “upper echelons of corporations where it’s boards, senior executives, and general counsels who care and who understand, and in some countries, they’re liable for those infractions.” While self-preservation might motivate some, “it’s a care for the company and not wanting to be exposed to that level of risk.” Even still, he argues, “ethics is doing the right thing when the wrong thing is possible or easy and legal.”

Our informants often couched the ultimate payoff of “doing ethics” as building a better product due to the in-depth product reviews and iterations that ethics would appear to demand. For example, one informant who serves as an ethics consultant framed ethics in terms of expanding what counts as a functioning product: “Just because your code functions, by the way, does not mean your product works, right? There is absolutely a social responsibility to learn and understand what the implications of your product will be, like, there is no skirting around that.... There’s absolutely responsibility to be knowledgeable.” From the inside, “ethics” is the creation of *better alignments* to these values, resulting in *smarter, safer, and fairer* products (Greene, Hoffman, and Stark 2019).

While the engineer wants to see more bottom-up engagement with ethics from her fellow engineers, the management consulting leader sees the future of ethics in the tech industry as the implementation of processes and protocols to eliminate organizational risk across a broad range of domains, from reputational risk to liability, as well as a hedge against being seen as inattentive to these risks (Power 1997). At the same time, they share the view that the practice of ethics would require financial or organizational sacrifice. These two ethics stakeholders come from different organizational vantage points, but they both also consistently refer to a set of cultural logics that are widely viewed as normative, if not desirable, in the tech industry.

THREE DOMINANT TECH INDUSTRY LOGICS

Tech firms are characterized by core logics that shape their organizations and are used to legitimate their power post hoc. These logics are mutually reinforcing and are therefore difficult to cleanly separate from

each other analytically. Put succinctly, Silicon Valley logics hold that trenchant social problems can be addressed through innovative technical solutions (Winner 2004; Segal 2005) developed by those with the most aptitude and creative energy (Neff, Wissinger, and Zukin 2005), and that an unencumbered market will recognize, reward, and disseminate the best solutions (Harvey 2005; Poon 2016). To be sure, there are other logics at play in the tech industry (Turner 2006), but these three interlocking components comprise a set of logics that justify both their own existence and the majority of activities that characterize the tech industry. These logics underwrite business as usual at the same time that they are implicated in many industry approaches to “doing ethics.” Throughout our interviews, ethics owners struggled with the dynamics of technological “solutionism,” meritocracy, and market fundamentalism, and all three consistently reappear as factors in their organizational calculus.

Meritocracy

An engineer we interviewed came right out and said it: “The tech industry claims to be a meritocracy. It is not.” This degree of reflexivity was rare in our interviews, however. Despite the origins of the term “meritocracy” in speculative fiction as a satirical indictment of British society for its retreat from liberal democratic values (Young 1958), meritocracy as a model of governance has unironically come to serve as a post hoc justification for inequality in society (Khan 2011). Unlike the aristocratic hierarchies that preceded it, meritocracy provides an ideological explanation for unequal distributions of wealth and power as arising from differences in individual abilities. Such differences in ability are often naturalized or otherwise reified while at the same time obscuring the power- and privilege-laden social structures that produce and perpetuate such inequalities (Ho 2009).

Meritocracy has become firmly entrenched in neoliberal conceptualizations of the modern subject as autonomous and responsible for perpetual self-improvement (Marwick 2013; Rose 1996), and signs of this capacity for self-improvement have acted as a kind of social and cultural capital that allows those with “merit” to have an outsized influence on public affairs. The tech industry has long been held up as a paragon of meritocratic achievement, in which its outsized economic and cultural power has been closely coupled with the technical and entrepreneurial skills needed to build and market products (Saxenian and Goldstein 1994). Meritocratic belief in its own abilities is a founding myth of Silicon Valley that animates the thinking of ethics owners we interviewed, who spoke of the imperative to “hire the best people” from “top schools” and to richly reward those who “are exceptionally skilled.” From this, a “can-do attitude” insinuates that those who work in the tech industry are sufficient to whatever task is presented to them, including the task of “doing ethics.” Meritocracy treats success as an index of an underlying quality that confers abilities that can be unreflectively applied to other domains. As one ethics owner explained, “there are really good people working in all of these companies who try to do the right thing.” Given the meritocratic thinking that underpins much of the tech industry, it is not surprising that many within the industry position themselves as the actors best situated to address the ethical challenges that have arisen in the last few years. In our interviews, one way this manifested was in positioning those within the industry to use their personal judgement as an instrument of moral action by “grappling with the hard questions on the ground,” trusting engineers both to discern and to evaluate the ethical stakes of their products. Certainly, engineers’ technical abilities came into play; they were sometimes presented as the people in the organization best positioned to evaluate whether a hypothetical harm was realistically possible given the technical capabilities of the particular application and, by implication, were free to dismiss the concern as unrealistic. But engineers are also seen as the locus of a company’s ethical sensibility; companies say they make very clear that “there are certain types of work we want to engage in [and] that was something that was reflected in decisions that employees came to the company” having made.

At the same time, many use the logic of meritocracy to dismiss critique or regulation. Notably, members of government who lack a grasp of fine technical details become easy targets. This is particularly salient in discussions around legislative or regulatory approaches to ethics. It is a routine experience at

“ethics” events and workshops in Silicon Valley to hear ethics framed as a form of self-regulation necessary to stave off increased governmental regulation. Meritocratic logics also manifested in ethics interventions that rely on personal accountability. Codes of ethics, statements of principle, checklists, and ethics trainings are oriented toward enabling engineers within the company to make “good” or “smart” decisions. Such tools center the role of the technologists as the locus of ethical agency, privileging their perspectives on potential harms produced by the products they develop and relying on their inherent goodness, as buttressed by meritocratic logics, to exhaustively address ethical issues. Under these conditions, when a problem emerges, blame can be placed on individual failure rather than institutional problems. This can create the conditions for tech workers to be what Elish (2016) calls “liability sponges.”

This is not to impugn the motives or ethical sensibility of those inside such companies; there is no reason to doubt that they seek to do good, ethical work, even when they make mistakes. Rather, the intention here is to suggest that the perspective tech workers have on broader societal problems is, at best, partial, as is their individual power within a corporate hierarchy.

Technological Solutionism

The partial perspective tech workers hold on broader social problems is a function of the cultural capital they hold and the economic capital with which they are invested. Both these forms of capital have been underwritten by the degree to which Silicon Valley tech firms have met with meteoric success over the past 25 years (Neff, Wissinger, and Zukin 2005; Saxenian and Goldstein 1994). The idea that technology can solve problems has been reinforced by the rewards the industry has reaped *for* producing technology that they believe *does* solve problems. Needless to say, critics quickly complicate that perspective, highlighting all the ways in which the tech industry’s “solutions” actually *cause* problems. Yet, even when members of the tech industry recognize their complicity in contributing to social problems, such as rising income inequality, they often respond by proposing technical solutions. It is therefore little wonder that ethical problems within the industry are often framed as challenges amenable to technological solutions. All too often, ethics is framed as a problem that can eventually be “solved” once and for all.

While some of our informants acknowledged that there may be no “silver bullet” for ethical problems, the search for technological solutions (Ball-Rokeach et al. 2004) that use “the super powers given to you by tech,” as one lead researcher put it, is strongly foregrounded in interviews and public discourse. Stated desires for toolkits and checklists posit that a technological solution is possible, even if it is yet to be fully developed. In some accounts, this gives the impression that ethics problems arise from imperfect technical solutions, which implies the inverse: ethical products will be better products. As a senior management consultant explained, “you’re not going to have an issue with the AI if your data is all good and you’ve curated that appropriately ... checked for fairness ... mitigated against negative bias ... and carried metadata along so that you can revisit decisions. ... And the problem exists when you don’t do those things.”

Technological solutionism contributes to an optimistic search for best practices—the optimal set of checklists, procedures, or evaluative metrics that will ensure an ethical product. Best practices have long been an important part of technological development work, ensuring computer code is clean and legible enough to be upgraded and debugged, and that hardware is manufactured to operate within the required performance specifications. This optimism is counterweighted by a concern that, when posed as a technical question, ethics becomes “intractable, like it’s too big of a problem to tackle,” which hampers collaboration between ethics owners and other technologists. Nevertheless, the framing of ethics as a technical problem persists. Describing ethics problems as best-practices problems centers ethics in the practices of technologists, not in social worlds they develop technical systems for and within.

Market Fundamentalism

Market logics profoundly shape the discourse around ethics in Silicon Valley and beyond (Duff 2016; Saxenian and Goldstein 1994). Ethics owners sometimes give voice to a cynical feeling that “[market] success trumps ethics,” and that in the absence of explicit prohibitions, companies might engage in deeply unethical practices so long as they are profitable. This cynicism grounds many of the calls for meaningful government regulation of the tech industry, as legal risk is a familiar constraint for tech companies. In the absence of regulations, ethics owners articulate a pressure to implement ethics practices that do not negatively affect companies’ bottom lines. As a senior leader in a research division explained, this “means that the system that you create has to be something that people feel adds value and is not a massive roadblock that adds no value, because if it is a roadblock that has no value, people literally won’t do it, because they don’t have to.” Market logics also exert a dampening pressure on ethics initiatives across the industry, as companies mirror each other’s approaches, sometimes out of fear of losing market share. As an executive explains, “if we play by these rules that kind of don’t even exist, then we’re at a disadvantage.”

This idea that “the bottom line speaks” was articulated in many different ways. A former social-media executive was lucid on how the demand side of the consumer market motivates corporate decision-making. “If people actually stop using certain services and/ or showed more preference for other services that didn’t have the same [design flaws], you can be sure companies will be responsive.” This reasoning was given as an explanation for why smaller or newer companies can’t be expected to develop their own sets of practices around ethics, as they are not profitable enough yet to be able to spare the resources on a capital-intensive project like “doing ethics.” This way of thinking also spurred a desire for durable processes and technical solutions around ethics. Once developed, they could be adopted by less well-capitalized companies, so that even small start-ups could “do ethics,” so long as the overhead costs for developing tools were carried elsewhere.

Given the pervasive market logics that saturate the tech industry, ethics owners often constrain their own capacity to effect change within the narrow remit of what “the market” might allow. Even the best-intentioned and most principled corporate officers, we heard, are faced with a tension between any number of ethical considerations and the fiduciary duty they hold to maximize shareholder value, fiduciary duty itself being understood as an ethical value (Lazonick and O’Sullivan 2000). While the market is sometimes seen as the hard and fast limit of what is possible in corporate approaches to ethics, some advocate for approaches aimed at reshaping the market to allow some greater degree of ethical intervention, including, for example, reforming regulation and oversight.

TECH ETHICS PITFALLS

The organizational logics of the tech industry are intersecting and mutually reinforcing. Meritocracy provides faith in the suitability of technological solutions to hard problems, while market success reinforces and constrains this correctness and further burnishes the credentials—and bank accounts—of tech workers. Many of the most common organizational approaches to addressing ethical concerns are clearly designed to fit within these logics. This is, in part, well-reasoned. By framing ethics as a difficult but tractable technological problem amenable to familiar approaches, ethics owners are able to enroll the technical and managerial experts they feel they need as full participants in the project of “doing ethics.” These approaches include ethics checklists and ethical project management frameworks, coding packages that evaluate algorithmic bias, and case study modules that resemble online learning styles with which software engineers are often familiar. Other proposed approaches are borrowed from adjacent fields, including “red teaming,”² statements of principles, and various codes of ethics, including a “Hippocratic oath for data scientists.” However, building a solution in the same mold that was used to build the problem is itself a mode of failure.

We heard ethics owners hedge against the notion that routine software development methods would yield adequate results (e.g., Johnson 2018). When ethics owners turn towards doing ethics in ways that do not adequately challenge these logics, their approach risks a premature return to the existing logics of tech development. This return to the everyday is a way of resolving the moments of public moral breakdown, when the ethical character of the firm's behavior is held up for inspection. In instantiating ethics through reproducible processes like checklists or programming packages, these efforts may manage only to offer the illusion of completion, and in doing so imply that ethics "has been done." Rather, it is precisely in these everyday practices that meaningful ethics are located, and robust anticipation of social harms could be contemplated. If these sensibilities are not refined but instead are prematurely foreclosed by technical interventions, then the problems that arise from the tech industry's way of doing ethics may persist. While there is a range of pitfalls that can unfold because of tech industry logics, we highlight two: normalizing ethical mishaps, and blinkered isomorphism.

Normalizing Ethical Mishaps

As tech companies start to envision procedures to operationalize ethics, they do so according to their own internal logics. As outlined by the senior leader in management consulting, one common way of approaching ethics is through legal or business risk accounting. Such an approach tasks senior leader with assessing the risk to the corporation in terms of legal, financial, or reputational risk (Olitzy and Benjamin 2001). Risk accounting describes risk and determines its acceptable levels, defining the situation in ways that let leaders manage potential consequences within a range of acceptable costs to the firm. Another approach is more deeply rooted in engineering and security practices. One of our informants noted,

you almost have to have the mentality of a red team that would look at this product and say, "How can somebody abuse this product?" Because it's so easy for people who are in that invention business to only look at the positive and say, like, "This is going to be so awesome. This is going to be so great." And then either forget or ignore that, hey, somebody's going to potentially take this product and use it for something that's really not good.

These practices resemble what Vaughan (1996) identified as "the normalization of deviance." Examining the communications records between engineers and managers working at NASA in the leadup to the *Challenger* shuttle disaster, she concludes that the horrific outcome was actually a byproduct of organizational actors responsibly and rationally developing, standardizing, and implementing risk procedures that, over time, produced significant vulnerabilities. In her words, "engineers and managers together developed a definition of the situation that allowed them to carry on as if nothing was wrong when they continually faced evidence that something *was* wrong" (1996). Far from being a one-off high-risk situation, the *Challenger* disaster was simply one moderate-risk event too many.

In Silicon Valley, a common motto is "fail fast, fail often." The industry rewards breaking rules and ignoring guardrails, normalizing deviance that is both similar to and different from what Vaughan described. While NASA has a clear framework for intolerable risks (e.g., loss of life, equipment, mission), tech companies have no formal boundary on which institutions should not be broken. Without such consideration, tech companies are creating the structural conditions to normalize ethical transgressions.

Without addressing the underlying logics to responsibly structure an organization to identify and guard against ethical failure, the mere implementation of ethical procedures may backfire, even when those processes were implemented to prevent that outcome (Pernell, Jung, and Dobbin 2017). Moreover, if ethics continues to be seen as something to implement rather than something to design organizations around, "doing ethics" may become a performance of procedure rather than an enactment of responsible values.

Blinkered Isomorphism

As tech companies begin to implement ethics practices across their organizations, there is pressure to share “best practices.” Yet, like journalists and scholars, they often focus on what *not* to do, highlighting public examples of corporate decisions that are now, in hindsight, deemed unethical. Less visible are the cases where employees and internal decisionmaking processes avoid the launch of products or features that might cause social harm. Companies rarely speak of products they chose not to launch. They also don’t address biases that their internal research identifies in their products until those products are publicly challenged (e.g., Buolamwini and Gebru 2018).

Collectively, with an eye towards avoiding risk, companies tend to steer their decisionmaking to respond to, and ideally avoid, public calamities. They are far less likely to share or learn from others’ successful actions. In this process, they create a form of “blinkerism.” Like a pony wearing blinkers, tech companies reproduce the same ethical blind spots across the entire industry through a process rooted in what DiMaggio and Powell (1983) call “institutional isomorphism.” This concept explains how companies and sectors converge on similar structures. In the context of tech companies, this can become “innovation through imitation” (Caplan and Boyd 2018). While institutional isomorphism and information asymmetries produce pitfalls for a range of organizations, in the context of ethics, tech companies risk mimicking each other against perceived extreme situations and creating blind spots around everyday ethical failings.

The isomorphic dynamic surrounding technology and ethics is shaped by corporate performance of ethics and public shaming of failure rather than direct exchange or interaction. Performing, or even showing off, the seriousness with which a company takes ethics becomes a more important sign of ethical practices than real changes to a product. As one engineer put it, “people want big dramatic responses.” *Performing* ethics, then, becomes a crucial component of doing ethics in tech. While an ethics management consultant admits that this can look “like ethics-washing companies” (akin to “greenwashing,” the act of superficially making environmentally unsustainable practices appear “green”), another engineer admits that in some cases “the appearance of effort matters more than results.” These performances come in many varieties, from the release of white papers and blog posts that proclaim companies’ searches for best practices, to corporate reorganization that promotes or creates a new ethics initiative.

Seen cynically, these performances produce what one engineer called a “cargo cult” mentality, a kind of magical thinking in which performing dramatic rituals is thought to produce material rewards, even if it also signifies a longing for the “moral, social, and cultural order in which these objects figure” (Crapanzano 2003). The more pragmatic consequence of this blinkerism is that companies are reshaping themselves in relation to an incomplete picture of the ethical landscape that others in their industry have traversed. They are learning to speak and perform ethics rather than make the structural changes necessary to achieve the social values underpinning the ethical fault lines that exist. This also presents a problem when approaches to ethics scale across enterprises. Isomorphism has been understood as a mechanism by which companies structure themselves in relationship to each other. But if companies are only publicizing their success while downplaying their failures, then there are few lessons to be learned for others, and an anemic set of tools may be the only approaches that circulate.

CONCLUSION

By talking with people who are at the forefront of thinking through ethics from within the technology sector, we found that the commitment to ethics is in tension with—and at risk of being absorbed within—broader and longer-standing industry commitments to meritocracy, technological solutionism, and market fundamentalism. Rather than challenging those commitments, the operationalization of ethics can work to uphold and affirm those other logics. Given this, the stakeholders that we interviewed had different perspectives on what should be done to make the most of the constraints that they collectively tolerated as

the cost of doing business in the tech industry. They admitted that the processes they were working to implement were necessary, if partial: “it’s got some flaws, but you know, at least it’s something.” Yet their tacit acceptance of at least one of these three core logics raises a significant question concerning ethics and tech: Can the tech industry be ethical if it does not call into question these underlying logics?

This question is increasingly prominent in critical data studies even as it appears that our work has at last found concrete footholds inside the tech industry, with scholars and critics demonstrating ambivalence about what “tech ethics” even means. D’Ignazio and Klein (2019) argue that the very use of the term “ethics” by the tech sector upholds what Ruha Benjamin (2019) describes as “imagined objectivity” because this term “locate[s] the source of the problem in individuals or technical systems.” Discomfited more specifically by the implementation of codes of ethics, Greene et al. (2019) provocatively note, “This presents a new problem for sociotechnical scholars used to being ignored: What if, instead of being brushed aside, our critiques are being heard but transformed into something we might not recognize?”

At the 2018 AI Now conference’s public symposium, human rights legal scholar Phillip Alston half-jokingly said from the stage, “I want to strangle ethics,” noting that ethics is “open-ended,” “undefined and unaccountable” in comparison to human and social rights frameworks (Alston 2018). Similarly, at the 2019 ACM Conference on Fairness, Accountability, and Transparency, the public feedback session was punctuated by members of the research community noting that “fairness,” “accountability,” “transparency,” and “ethics” can operate as “weasel words” that act as empty gestures, and that “justice, liberation, and rights” are frameworks for material improvement of human lives. While we share this ambivalent stance toward ethics as a programmatic goal, the oddity of this moment should not go unremarked: just when traction becomes possible, we are called upon to question whether we are climbing the right hill at all.

Our research shows that ethics owners live inside this same tension. For better or worse, their collective goal is not to stop the technology industry. Although they are all engaged in some form of critique of their own industry, they are also enmeshed in organizational cultures that reward metric-oriented and fast-paced work with greater resources, thereby ratcheting up the ability to fit in and ratcheting down the capacity to object. In a situation like that—just as it is with current critical data studies scholarship—it is necessarily challenging to distinguish between success and failure. Thus, the work of ethics consists of working through, and not resolving, these tensions.

While it is too early to know if the nascent efforts to effect change in tech companies will actually produce more ethical organizations, we are concerned that the organizational, cultural, and structural conditions inhibit such lofty goals. If ethics is simply absorbed within the logics of market fundamentalism, meritocracy, and technological solutionism, it is unlikely that the tech sector will be able to offer a meaningful response to the desire for a more just and values-driven tech ecosystem. This is a fate we would prefer to avoid. The question at hand for both critical data scholars and ethics owners inside these companies is how to find a route to an improved mode of “doing ethics” that institutionalizes a more open, just, and critical everyday practice.

ACKNOWLEDGMENTS

The authors wish to thank our informants for generously sharing thoughts and time, and our colleagues who took the time to read this work in draft form: Ryan Hagen, Alondra Nelson, Ben Roome, Janet Vertesi, Karen Strassler, and the Raw Materials Seminar group at Data & Society. We also wish to thank the NSF (awards #1704369 and #1633400) and the Wenner Gren Foundation for supporting portions of this work.

NOTES

1. We use “Silicon Valley” and the “tech industry” interchangeably to refer to the range of companies that are typically described by US-based public discourse as such. These companies are typically for-profit, consumer-facing, California-based, and significantly capitalized either as public companies or via venture capital. *Owning Ethics* 473

2. A “red team” is an internal team whose sole function in the company is to imagine and execute ways of undermining the company’s technology.

REFERENCES

- Alphabet Inc. 2018. Form 10-K.
<https://www.sec.gov/Archives/edgar/data/1652044/000165204419000004/goog10-kq42018.htm>.
- Ananny, Mike. 2016. “Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness.” *Science, Technology, & Human Values* 41 (1): 93–117.
- Ball-Rokeach, S. J., M. Sturken, and D. Thomas, eds. 2004. *Technological Visions: The Hopes and Fears that Shape New Technologie*. Philadelphia: Temple U. Press.
- Benjamin, Ruha. 2019. *Race after Technology*. Boston: Polity Press.
- Buolamwini, Joy and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” In *Proceedings of Machine Learning Research* 81: 15. New York.
- Caplan, Robyn and danah boyd. 2018. “Isomorphism through Algorithms: Institutional Dependencies in the Case of Facebook.” *Big Data & Society* 5 (1).
- Crapanzano, Vincent. 2003. “Reflections on Hope as a Category of Social and Psychological Analysis.” *Cultural Anthropology* 18 (1): 3–32.
- Das, Veena. 2016. “What Does Ordinary Ethics Look Like?” In *Four Lectures on Ethics: Anthropological Perspectives* 73. Chicago: Hau Books.
- Dave, Naisargi. 2012. *Queer Activism in India: A Story in the Anthropology of Ethics*. Durham, NC: Duke U. Press.
- D’Ignazio, Catherine and Lauren Klein. 2019. *Data Feminism*. MIT Open Press.
<https://bookbook.pubpub.org/pub/data-feminism>.
- DiMaggio, Paul and Walter Powell. 1983. “The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields.” *American Sociological Review* 48 (2): 147.
- Duff, Alistair S. 2016. “Rating the Revolution: Silicon Valley in Normative Perspective.” *Information, Communication & Society* 19 (11): 1605–21.
- Elish, M. C. 2019. “Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction.” *Engaging Science, Technology, and Society* (5): 40–60.
- Fiesler, Casey. 2018. “Tech Ethics Curricula: A Collection of Syllabi.” *Medium*, July 5.
<https://medium.com/@cfiesler/tech-ethics-curriculaa-collection-of-syllabi-3eedfb76be18>.
- Fleischman, Gary, Eric Johnson, Sean Valentine, and Kenton Walker. 2017. “Ethics versus Outcomes: Managerial Responses to Incentive-Driven and Goal-Induced Employee Behavior.” SSRN Scholarly Paper ID 3032078.
<https://papers.ssrn.com/abstract=3032078>.
- Greene, Daniel, Anna Lauren Hoffman, and Luke Stark. 2019. “Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning” 10. Hawaii International Conference on System Sciences (HICSS), Maui.
- Harvey, David. 2005. *A Brief History of Neoliberalism*. London: Oxford U. Press.
- Ho, Karen. 2009. *Liquidated: An Ethnography of Wall Street*. Durham, NC: Duke U. Press.
- Johnson, Nick. 2018. “Why Ethical AI Is a Mindset, Not a Checklist.” *Salesforce Blog*. Nov. 2.
<https://www.salesforce.com/blog/2018/11/ethical-ai-is-a-mindset-not-a-checklist.html>.
- Keane, W. (2017). *Ethical Life: Its Natural and Social Histories*. Princeton, NJ: Princeton U. Press.
- Khan, Shamus Rahman. 2011. *Privilege*. Princeton, NJ: Princeton U. Press.
- Laidlaw, James. 2002. “For an Anthropology of Ethics and Freedom.” *Journal of the Royal Anthropological Institute* 8 (2): 311–32.

- Lambek, Michael. 2010. *Ordinary Ethics*. New York: Fordham U. Press.
- Lazonick, William and Mary O'Sullivan. 2000. "Maximizing Shareholder Value: A New Ideology for Corporate Governance." *Economy and Society* 29 (1): 13–35.
- Mahmood, Saba. 2012. "Ethics and Piety." In *A Companion to Moral Anthropology*, edited by Didier Fassin. Malden, MA: John Wiley & Sons.
- Marwick, Alice. 2013. *Status Update: Celebrity, Publicity, and Branding in the Social Media Age*. New Haven, CT: Yale U. Press.
- Neff, Gina, Elizabeth Wissinger, and Sharon Zukin. 2005. "Entrepreneurial Labor among Cultural Producers: 'Cool' Jobs in 'Hot' Industries." *Social Semiotics* 15 (3): 307–34.
- Orlitzky, Marc and John Benjamin. 2001. "Corporate Social Performance and Firm Risk: A Meta-Analytic Review." *Business & Society* 40 (4): 369–96.
- Pernell, Kim, Jiwook Jung, and Frank Dobbin. 2017. "The Hazards of Expert Control: Chief Risk Officers and Risky Derivatives." *American Sociological Review* 82 (3) 511–41.
- Poon, Martha. 2016. "Corporate Capitalism and the Growing Power of Big Data: Review Essay." *Science, Technology, & Human Values* 41 (6): 1088–108.
- Power, Michael. 1997. *The Audit Society*. New York: Oxford U. Press.
- Rose, Nikolas. 1996. *Inventing Our Selves*. New York: Cambridge U. Press.
- Saxenian, AnnaLee and Karl Goldstein. 1994. *The Software Industry in Northern Alameda County: Development Patterns and Prospects*. Berkeley: U. California Berkeley, Institute of Urban and Regional Development.
- Segal, Howard. 2005. *Technological Utopianism in American Culture*. Syracuse, NY: Syracuse U. Press.
- Simonite, Tom. 2019. "Google and Microsoft Warn that AI May Do Dumb Things." *WIRED*. Feb. 11. <https://www.wired.com/story/googlemicrosoft-warn-ai-may-do-dumb-things/>.
- Sims, Ronald R. 1991. "The Institutionalization of Organizational Ethics." *Journal of Business Ethics* 10 (7): 493–506.
- Trevino, Linda Klebe. 1986. "Ethical Decision Making in Organizations: A Person-Situation Interactionist Model." *Academy of Management Review* 11 (3): 601–17.
- Turner, Fred. 2006. *From Counterculture to Cyberculture*. Chicago: U. Chicago Press.
- Vaughan, Diane. 1996. *The Challenger Launch Decision*. Chicago: U. Chicago Press.
- Winner, Langdon. 2004. "Sow's Ears from Silk Purses: The Strange Alchemy of Technological Visionaries." In Sturken and Thomas, 2004, 34–47.
- Young, Michael. 1958. *The Rise of the Meritocracy 1870-2033*. Bristol, UK: Thames and Hudson.
- Zigon, Jarrett. 2007. "Moral Breakdown and the Ethical Demand: A Theoretical Framework for an Anthropology of Moralities." *Anthropological Theory* 7 (2): 131–50.
- . 2014. "An Ethics of Dwelling and a Politics of World-Building: A Critical Response to Ordinary Ethics." *Journal of the Royal Anthropological Institute* 20 (4): 746–64.