In the Walled Garden: Challenges and Opportunities for Research on the Practices of the AI Tech Industry

Morgan Klaus Scheuerman
morgan.scheuerman@colorado.edu
University of Colorado Boulder
Boulder, Colorado, USA

ABSTRACT

Research on technology companies and their workers can externalize otherwise invisible and tacit workplace approaches, identify organizational constraints to creating more ethical AI systems, help ground interventions in real-world organizational realities, and result in the co-creation of better business practices for organizations. However, getting access to technology companies is difficult for external researchers. In this paper, I draw from insights gained by conducting research on and with industry professionals. I present four challenges when conducting industry-focused research on responsible AI. I also present methods I used to navigate each challenge. Finally, I highlight opportunities for the tech industry to lower the barriers to external research. This work aims to share ways of navigating methodological challenges and encourage better transparency in the tech industry.

CCS CONCEPTS

• **Social and professional topics** → **Computing industry**; • **Human-centered computing** → *Empirical studies in collaborative and social computing*; • **Computing methodologies** → Artificial intelligence.

KEYWORDS

Human subjects research, methods, big tech, tech industry, transparency, work studies, AI ethics

1 INTRODUCTION

The diligent work of researchers, policymakers, and public stakeholders have contributed great strides towards increasing the fairness, accountability, and transparency of commercial artificial intelligence applications. For example, Buolamwini and Gebru's prominent GenderShades work led numerous companies to audit and update their computer vision models [64]; Google, and soon other

companies, similarly began removing gender labels from their vision models following academic work critiquing its exclusion of trans and non-binary people [42]. Beyond direct lines between specific work and product changes, broader critiques around ethical AI implementation can still permeate the culture of companies building AI systems, leading to corporate declarations of dedication to responsibility and ethics[59, 83].

However, audits, studies, and investigations of publicly available data can neither fully uncover ethical issues in existing AI applications nor the practices and procedures that lead to problematic outcomes, whether normative or quantifiable. While internally conducted studies of AI systems can address shortcomings of external investigations, they rarely result in meaningful public transparency. Thus, external third-party[1] researchers are increasingly examining the practices of industry to identify what major barriers stand in the way to implementing ethical practices in corporate settings (e.g., [37, 65, 85]). External research can offer new perspectives from which to understand, and intervene, in the tech sector.

Given the opacity and monopolizing power private companies have on the development and deployment of AI systems, the work of external researchers untethered to these companies is crucial to piecing together a picture of the general landscape of the internal workings of key decision makers. However, getting meaningful external access to these companies—their employees, products, data, or work sites—is difficult. As big tech has evolved from an ethos of innovation to a model of increasingly centralized data, patents, and knowledge [44, 67, 84, 90], the walls to accessing tech companies as research sites have gotten increasingly higher. Higher barriers for external researchers do not simply mean that researchers cannot publish novel insights into company practices. It means that the broader academic community, policymakers, and the public are unable to obtain information from companies outside of what those companies publish, which is not dictated by any meaningful regulation to ensure transparency. It also means that academic considerations and policy regulations might not represent reality within the walls of tech companies. How then can the broader community identify significant product biases, unfair or problematic practices, or hold companies accountable?

In this paper, I present insights into the major challenges that I encountered when conducting and publishing human subjects research on industrial AI practices. The social practices behind industrial AI development is particularly pertinent given the number of isolated parties involved in developing and using AI, from data workers, full-time tech workers, clients, and regulators. In

---

[1]In this article, I use the terms first-party (internally employed tech workers), second-party (contractors or collaborators under NDA), and third-party (entirely external independent researchers) as defined by Costanza-Chock et al. in [16]

this work, I draw from an examination of numerous third-party and second-party [16] research experiences with the tech industry. Specifically, this paper: (1) documents four challenges to conducting research on the AI practices of industry practitioners; (2) provides approaches for navigating these challenges, while actively acknowledging the transparency tradeoffs in doing so; and (3) highlights opportunities for the tech industry to mitigate each of these challenges, ideally, without making transparency tradeoffs. I emphasize that the tech industry itself would also benefit from improving access and transparency regarding external research as well.

The goal of this work is not to universally ascribe a taxonomy of challenges, navigations, and opportunities. Certainly, there are many more challenges and opportunities for dealing with them. Instead, I make explicit some of the otherwise implicit challenges faced by external researchers when conducting work on the social practices in the tech industry. Further, this paper offers approaches to navigating these challenges that researchers might adopt and adapt. Finally, this paper highlights opportunities for the tech industry itself to actively consider how to shift current policies and practices to better engender research transparency during this time of heightened scrutiny of the opacity of tech companies.

2 BACKGROUND

In response to the rapid growth of machine learning and artificial intelligence, researchers have been at the forefront of identifying, measuring, and attempting to mitigate biases in AI systems. Interdisciplinary researchers have conducted audits of commercial models (e.g., [2, 58, 64, 75]) and community benchmarks (e.g., [5, 6, 63]); developed taxonomies, frameworks, and toolkits (e.g., [11, 17, 26, 33, 41, 46]); presented sociotechnical analyses of the values present in AI and how they can lead to harms (e.g., [21, 45, 74]); and assessed user experiences and expectations of AI ([15, 28, 35, 48, 79]). Largely, much of this work has been accomplished through engaging with publicly available artifacts (models, datasets, publications, documented AI incidents (e.g., [25])).

Yet the most impactful AI systems continue to be developed rapidly inside company walls. Engagement with public-facing artifacts can only paint a partial picture of the current state of AI. Thus, researchers are increasingly focused on understanding current practices *within* the otherwise opaque organizations developing and implementing industrial AI. For example, prior work has found that workers attempting to implement human-centered practices at technology companies face barriers (e.g., [16, 32, 37, 65, 85]) and must engage in extra articulation work to bridge gaps, piggyback off of more incentivized institutional procedures, and otherwise resist practices they disagree with (e.g., [1, 9, 20, 68, 88]). Such work is important, not only because it reveals insights normally hidden within the walled garden that is the tech industry, but because it allows external stakeholders to ground their work, interventions, and advocacy in real, rather than perceived, practices and product issues.

However, absent from many of the successful research studies headed by external researchers are the methodological realities: work involving tech companies is difficult. Getting access to tech companies for external research, either as a third-party researcher

or a second-party collaborator [16] requires intense effort and numerous tradeoffs—particularly when that work is focused on ethics and fairness topics. Tech companies have largely shuddered access to external researchers (e.g., [10, 19]), even to their otherwise public-facing user data [27]. Even practices which allow for access to anonymized user data, such as "data clean rooms," [2] are vague and non-standardized. Tech workers have also increasingly described instances where research agendas focused on work more critical of company practices and products have been downplayed or silenced [50, 66, 76]). Beyond improving their PR, tech companies are losing out on valuable knowledge that is rooted in actual tangible practices (e.g., [30]). Rigorous, truthful, and reliable external research has the potential to benefit both the public and the company.

As Obermeyer stated about examining the otherwise unnamed healthcare algorithm in [58]: "Getting a handle on the algorithms that are live within an organization is critical" [51]. I would extend this to "getting a handle on the practices that are live within an organization is critical" as well. The notion of "auditing" AI can extend beyond developed artifacts to examine the processes and negotiations that go into developing them. However, understanding these social processes remains difficult for external researchers. The challenges present in conducting human subjects research on the AI industry remain invisible and implicit. In this work, I describe the numerous challenges that external researchers may face in "getting a handle" on the internal mechanisms of the tech industry and how the tech industry can increase opportunities for external research.

3 METHODOLOGY

In this work, I draw from experiences conducting responsible AI research on and within technology companies (e.g., [71–73]). The experiences summarized in this work span the last five years, in which I conducted both semi-structured interviews and ethnographic work [80]. Ethnographic work is gaining momentum in FAccT and responsible AI spaces (e.g., [52]); with it comes specific barriers that have been underdocumented. In this work, I adopt methods for "studying up" [4, 54, 55], focusing on analyzing my ethnographic data from various projects from a methodological lens that centers the power of tech companies over independent research practices. The barriers I present reflect positions I have held as both a third-party researcher, totally independent from industry, and as a second-party researcher, collaborating with industry researchers and holding temporary contracts with industry myself. While there are also challenges to conducting industry research from a first-party perspective (as a full-time industry employee), I present barriers specific to my experience in both third- and second-party roles.

In my capacity as a third-party researcher, I reference barriers from three empirical studies focused on practices and perspectives of industry stakeholders––those employed by technology companies as either traditionally employed full-time workers (e.g., product managers, engineers, researchers) or as contingent workers (e.g., data annotators, data collectors, content moderators). Throughout

---

[2]Environments where first-party user data is anonymized and securely shared by platforms, largely aimed at advertisers and marketers looking to buy consumer data [77]

these three studies, I have spoken to a total of 86 research participants who work in tech and conducted over 400 hours of interviews and observations. I synthesize these findings with reflections on my experiences as a second-party researcher working within companies and collaborating with researchers internally employed at companies. To analyze my data, I performed a thematic analysis [13] of both field notes and interview transcripts focused on identifying methodological challenges and barriers to industry research. Beyond inductively analyzing my research documents and field notes for larger themes, I also discussed these themes with colleagues in industry and academia, to assess both accuracy and potential solutions.

4 CHALLENGE 1: IDENTIFYING TECH WORKERS FOR THIRD-PARTY RESEARCH CAN BE DIFFICULT

The recruitment of tech workers to participate in external research studies, like semi-structured interviews, is difficult. Unlike other populations, tech workers are not easily recruited using typical recruitment methodologies, like surveys posted to social media or community spaces [56, 69]. I had tried traditional forms of participant recruitment for my studies. Using social media websites, like Twitter and Blind, was not successful. For example, I had tried recruitment by having industry colleagues post a call on Blind[3], neither of which were successful and largely resulted in critique and skepticism of my intentions (see Challenge 2). I found that recruiting tech workers for research studies is often more easily accomplished through the direct recruitment of individuals, either through introductions from tech insiders or by "cold calling" (directly reaching out).

However, To recruit using this method, the researcher must be able to identify their potential research subjects. Many tech workers are not public facing. In particular, those in more technical roles, like data scientists and software engineers, had less clear web presence than those in research or C-level roles[4]. Others involved in the tech industry, such as contingent workers contractually employed by tech companies, were entirely invisible. Contingent workers contracted by companies largely cannot disclose who they conduct work for. When certain tech worker populations are easier to reach (e.g., research scientists), insights into practices are limited towards that population's perspective.

The issue of identifying potential participants to recruit is further compounded when attempting to target specific technical domains. For example, in two of my studies, I was specifically seeking tech workers in a relatively narrow subfield of machine learning: computer vision. The more specific the populations researchers are attempting to reach, the more difficult the process of identifying potential research subjects becomes. It also becomes increasingly important, given many participants I recruited had highly limited knowledge about the history of computer vision products at their company. Being able to identify multiple workers who have experience with the same product can help fill knowledge gaps and point researchers toward colleagues who could provide further insight.

4.1 Navigating Challenge 1: How Researchers Can Better Recruit Relevant Tech Worker Participants

The inaccessibility of tech workers can often be attributed to a lack of visibility, often because external researchers may be highly divorced from more product-centric roles. Difficulty identifying participants to recruit was particularly salient for those in technical and product roles, like software engineers, data scientists, and project managers. In navigating this challenge, researchers should consider conducting their work with the people they can get access to, rather than aiming for "perfect" access.

4.1.1 Build relationships in the tech industry. Leveraging insider relationships was the most successful means of recruitment. I developed insider relationships with industry employees through research collaborations and consulting. Thus, mutual connections trusted me more as a researcher due to my insider relationships with others at the company. In one project, I had access to a company email. Using a company email was also helpful in facilitating trust between myself and those I was trying to recruit. I also gained access to a data production company as a field site by connecting with an acquaintance I had made at a workshop on LinkedIn. Researchers should consider attending webinars, public talks, conferences, and local events to build relationships with industry stakeholders (e.g., [1]). Researchers should focus on building relationships with workers in spaces they are most interested in studying (e.g., computer vision). However, there were many cases where individuals did not respond, even with their mutual connection's facilitation. This might point to fear of research participation (see Challenge 2).

4.1.2 Identify potential participants using digital tools. Given much of my research focus has been on computer vision, I located computer vision companies through search tools, like Google and LinkedIn. I was also able to identify individuals and their roles using LinkedIn's search tools and RocketReach[5]. To try to find contingent workers who conducted labeling for tech companies, I used freelancing platforms like Upwork to hire workers for interviews. Identifying the type of work employees do can still be difficult with just their company's name. Therefore, I often reached out to individual employees describing the focus of my study and requesting further information on their roles.

4.1.3 Budget more time for recruiting industry participants. Recruitment of industry participants can take time. I had to put much of my work on hold for a couple of years as I established more connections with industry. Even after I had established connections, iteratively recruiting tech worker participants for a single study took about a year. When conducting ethnographic work at a data production company, I spent a great deal of time building a relationship with the CEO. Given the difficulties of identifying and recruiting tech worker participants, researchers should budget more time for recruitment stages than they might in more user-centered studies.

---

[3]Blind is an anonymous forum-style social media website catered towards verified employees in the tech industry; https://www.teamblind.com/
[4]Senior business leaders with high-ranking executive titles (e.g., CEO, CFO, COO)

[5]RocketReach is a subscription-based database of professionals; https://rocketreach.co/

4.2 Opportunity 1: How Industry Can Break Down Walls Between Tech Workers and Researchers

Gaining access to participants and field sites, particularly for long-term studies like ethnographies, is understandably difficult [62, 80]. However, corporations have the opportunity to make it easier to connect external researchers and tech workers. Building connections with external researchers is beneficial to companies looking to transform research insights into practice, maintain relationships with academic research institutions, and build a more transparent brand. Below, I ideate some ways tech companies might operationalize easier access:

- Tech companies can be more transparent about their personnel. For example, companies can consider how best to list employees who are available for research contact. They might allow individuals to sign up for research contact. Lists of potential participants could be high-level or granular to teams or roles. Press relations experts could be trained to connect researchers with internal participants.
- Internal employees often recruit their own colleagues for research studies through official channels. Tech companies could build a platform for external researchers to host calls for participation.
- Tech companies might hold more regular events specifically focused on connecting researchers with opportunities for research collaborations.

There are still tradeoffs to these potentials, especially in terms of selection biases. Workers who would opt into these types of opportunities might have some personal reasons for doing so [24]. The company might also use these opportunities to promote certain worker perspectives or develop new mechanisms for moderating research communications (see Challenge 4). The research community might instead choose to accept the current challenges, rather than create more formalized procedures for recruitment.

5 CHALLENGE 2: THE NDA CAUSES ANXIETY FOR POTENTIAL PARTICIPANTS AND CAN CAUSE DIFFICULTY WITH PUBLICATION

Non-disclosure agreements (NDAs) are a challenge to navigate. NDAs were challenging in two ways. First, in instances where external researchers are conducting third-party research with tech workers, NDAs were a major source of anxiety and distrust among potential participants. The tech workers I spoke with were afraid of accidentally violating their own NDAs. Numerous participants mentioned their NDAs during interviews. For example, a developer advocate mentioned during our interview that she did not want to *"[get] in trouble for our NDA thing."* NDAs also remained consequential for participants even after they had left their companies. A participant who had left her company over ethical concerns stated that she had never talked about her experience with her former company. As she told me, *"The NDA is really the fear."* During early recruitment stages, I discovered from colleagues that internal discussions were concerned about discussing anything related to identity and AI. They were afraid they might accidentally open themselves up to legal liability with their employers. Similarly,

some participants who had agreed to participate backed out before the interview due to legal concerns surrounding fresh controversy at their companies.

Further, participants did not trust external researchers. Much like others have found [37, 82], tech workers also seemed to have an inherent distrust of external researchers. Some participants expressed concerns that the academic community was reactionary towards industry. Participants with a distrust of academics seemed to view academics and journalists similarly, assuming that both were interested in uncovering a *"clickbait headline"* to publish about tech companies. This distrust amidst a wave of articles covering AI ethics was salient among participants I talked to throughout my research. For example, the head of data science at a small company focused on AI interviewing described how she felt academics demonized industry practitioners:

> We're very aware of the knee-jerk reactions people have to what we do ... I know what it's like to be in academia, and sometimes you think, *'These people that are in business are just like greedy, moralless, soulless people and they don't think about the consequences or the nuances to what they're doing.'* But that's definitely not true.

To address their concerns about violating their own NDAs and distrusting external researchers, tech workers often sought to have external researchers sign an NDA. I learned that many tech workers had backchanneled about asking me to sign an NDA amongst their peers. A participant who had initially declined to participate, but later participated after I had built a relationship with her, informed me that my initial recruitment emails had caused a great deal of *"backchanneling"* about whether participation was too risky without having me sign an NDA:

> I'll summarize the back response to your email ... We had someone who said, *'It'd be fascinating to talk to them. I have a lot of thoughts, but I'm not sure how much I should or can say. I'd feel more comfortable if they could sign an NDA.'* Uhm. Then we have someone else who says, *'My general rule of thumb is that anything we've published externally is on the table for discussion. Anything we haven't should be kept abstract or high-level unless they're under NDA.'*

While not necessarily explicitly, workers hoped to ease anxieties about violating their own NDAs by putting legal liability of violations onto external researchers. However, most external researchers understand that being under NDA severely hinders the publication of research results with the scholarly community—the major goal of conducting such research in the first place. In cases where I was working as a second-party researcher and was under NDA, I faced challenges publishing the results. However, unlike first-party researchers employed directly by a company, external researchers have options for navigating NDAs when they must sign them—which I discuss in the next section.

5.1 Navigating Challenge 2: How Researchers can Build Trust with Tech Worker Participants and Publish Even Under NDA

Participants felt speaking to external researchers was a risk. Navigating the anxiety and distrust that tech workers had about participating in academic research meant assuaging their concerns about their NDAs and their distrust of external researchers. While the two solutions below mean research is still opaque, in the sense that specific company practices are still hidden from the public eye, they allow academic researchers to communicate key insights that can be generalized to other industry practices, products, and stakeholders. Given the current state of conducting external research on the tech industry, opacity about a field site is often a necessary tradeoff for transparency of knowledge.

5.1.1 Reassure tech worker participants by stating your goals and procedures upfront. Much like all forms of human subjects research, working with tech workers meant building trust and alleviating concerns they have to participate in research. Navigating the fear that tech workers had about participating in research meant engaging explicitly with common concerns among tech workers during recruitment and rapport building stages. I became more upfront about my intentions in my recruitment and often reassured them that I had no intention of publishing information to harm them or their companies. I would make sure to tell participants that I was interested in understanding their work practices, not in unearthing scandalous information about their companies. I also pointed them to the information about the anonymization process in consent forms and study information sheets. Being explicit about my intent helped participants to feel more comfortable talking with me and even air their grievances with academic researchers more openly. Researchers might also consider the pros and cons of allowing their research participants to "member check" any quotes or descriptions they might use in publications [14, 29, 34]. Finally, some institutions, like my own, have legal representatives who will review proposed research and certify when a study has minimal legal risk to participants.

5.1.2 Consider signing the NDA and then disseminating your research results carefully. In some cases, researchers might need to take on a second-party role and sign an NDA in order to conduct research on industrial practices or products. Having worked with numerous companies, I have signed several NDAs. Researchers still have options for publishing without violating their NDAs. While many times those working full-time at companies have to have their research approved by compliance review boards (see Challenge 4), researchers in temporary roles have some opportunities for bypassing corporate procedures. I conducted research within a company while under NDA, but later published the work with permission and guidance from the internal team members at the company. Tech worker colleagues can provide guidance on how best to avoid violating NDAs or corporate compliance review processes. Researchers in this situation should also realize that they may not have access to the data they collected forever—the data collected for the aforementioned project was eventually revoked by the company. Researchers who are not allowed to make anonymized copies of the data themselves might consider writing up anonymized notes about the project or prioritize writing up their research results quickly.

5.2 Opportunity 2: How Industry can Rethink their Approach to NDAs

While the distrust of academic researchers is largely on researchers themselves to alleviate, the fears that tech workers have about violating their NDAs is something that tech companies could address. While companies utilize NDAs to protect intellectual property and prevent issues of legal liability, they would also benefit from external insight into the perspectives and practices of their workers and teams. The common use of NDAs−−with their employees, with external researchers, and with temporarily embedded researchers−−leads to less transparent research publication. Here, I offer a few recommendations for tech companies to consider concerning NDAs:

- NDAs should not be used to silence tech workers from speaking broadly to external researchers or other public stakeholders about their perspectives and criticisms, observed practices, or specific topics (e.g., identity, fairness).
- The majority of external academic researchers are not requesting specific information on trade secrets or intellectual property. Yet, workers may sign NDAs with little understanding of what would violate them. Tech companies might provide easily accessible training or guidelines on navigating the NDA.
- In cases where a researcher is asked to sign an NDA to conduct research, the researcher and company should be able to come to an agreement about exception clauses so that useful research insights can still be communicated—even if compliance reviews are still involved (see Challenge 4).

Here, I openly acknowledge that companies will always prioritize protecting intellectual property over contributing to external research. There has been considerable criticism of NDAs for cloaking unethical business practices [7, 8, 23, 39, 61, 87]. Certainly, law and policy are the proper avenues for reconsidering the current practice of NDA use [3, 22, 43, 49]. Legally limiting the use of NDAs could both ease the anxieties of individual employees looking to discuss their experiences with researchers and repeal the negative consequences many external researchers see about signing NDAs to collaborate with industry.

6 CHALLENGE 3: TECH WORKERS LACK INCENTIVES TO ENGAGE WITH EXTERNAL RESEARCHERS

Many tech workers might not feel any incentive to participate in third-party research. There were no formal or informal mechanisms that incentivized full-time tech workers to interact with external researchers or otherwise provide transparent insight about their work to the broader public outside their company's doors. Participating in external research not only opens tech workers up to personal risk (see Challenge 2), but it does not contribute to their growth within the company.

Incentives within companies are largely driven by business needs first and foremost. As a user experience researcher at a large tech company explained, *"The incentive structure is just to create products*

and ship it." Even when workers were interested in participating, other formal and informal mechanisms at tech companies made it difficult or impossible for them to participate. Throughout my work with tech companies, numerous colleagues lamented that the research they wanted to do was not measurable or quantifiable in the form of Key Performance Indicators (KPIs). KPIs are largely measured in the form of "impact," which in many tech companies is tied directly to the development of a product, rather than contributing to broader knowledge outside of the company. Even those in more academic-style research roles could not set their own research agendas; research insights had to contribute in some form toward specific product or company goals.

Even when external researchers are welcomed through the doors of a company as collaborators, some participants I spoke with felt that external researchers were not interested in attending to their company's priorities. For example, a C-level executive at a small company described how his company often tried to involve academics in joint research endeavors. However, he said that *"some of the discussions fall apart because it's not the exact data they want ... or they don't want to modify their research agenda to fit the data that we have ... They want us to change our practice to fit [their research]."*

Even when companies valued second-party academic collaborations, experience working with researchers in the past led company representatives to view external researchers as unable to provide any direct benefit to the company. Industry stakeholders may find that the demands of external researchers don't benefit their goals or the labor they put in (an already common research ethics discussion when it comes to working with other communities [78]). The lack of perceived benefit might be further compounded by the pace at which industry moves in comparison to academia, as well as the approach in academia being more systematic and less focused on key product takeaways.

Given the combination of personal and legal risk, a lack of formal incentive, and the perception external researchers do not add value to industry work, those individual tech workers interested in participating in external research is often driven by personal values about the importance of broader knowledge dissemination. Such individuals might have to dedicate time and energy to navigating legal constraints, forgoing their own professional growth. Individuals without these values may also hold important knowledge, yet they are unlikely to participate in research. The lack of incentives provided by tech companies themselves also means that researchers must spend more time and resources attempting to create incentives for an already difficult to reach population.

6.1 Navigating Challenge 3: How Researchers can Incentivize Individuals and Companies with Research Insights

Currently, there are few incentives for tech workers to participate in third-party research, especially if tech workers are in product-centered roles where research outcomes are not considered a part of their KPIs at all. Even those researchers who conduct internal first-party research are incentivized to pursue research which has tangible impacts for the company, rather than basic research. As a result, the tech workers most likely to participate in third-party

academic research have a personal value bias—they likely believe in the ethos of knowledge sharing that academic research represents[6]. There is little that external researchers can do to change how a company rewards individual workers for participating in research. However, there are methods for incentivizing individual tech workers to participate in research. There also ways to provide useful and actionable insights to a company, which might otherwise improve relationships with external researchers.

6.1.1 Consider how best to incentivize individual tech workers to participate in third-party research. While a common approach in human subjects research is to financially compensate their participants, financial incentives may hold less value for participants in highly paid, highly valued professional roles, like many tech workers employed full-time at technology companies are. Beyond the lack of value, some companies do not even allow their workers to accept financial compensation for research participation. Certainly, I found that financially compensating workers in contingent roles worked as an incentive, as those workers tend to be low paid and suffer financial instability][31, 81, 86] in comparison to their full-time peers located in the Global North. Of course, providing adequate incentives might be particularly challenging for researchers from low resource institutions, from disciplinary fields that are not traditionally seen as valuable to tech companies, and who are focused on problems that are not of interest to the company or workers. If researchers hope to incentivize high-paid workers employed full-time at companies, they may have to think outside of the box of traditional human subjects research.

6.1.2 Share research insights back with participants in ways that are meaningful to their performance evaluations. Researchers can work with participants to understand how research can add value to them personally. For example, an engineer will likely not receive any form of recognition for participating in a study published at an academic venue. However, if the researcher offers to work with them to write a short white paper or offer an internal talk to their research team, they can help present the work as a collaboration that the company values in the form of KPIs.

6.1.3 Consider how to demonstrate the usefulness of tech workers' contributions when invited to collaborate second-party research within a company. While the researcher should not sacrifice all of their goals on the research project, they should negotiate with their industry partners to also address problems they care about, using data and practices the company already employs. They should discuss with their field site partners what outcomes would be most useful to them. For example, I have been working with a data production company to develop a training or webinar that will be disseminated to their data workers. Researchers might also consider delivering ongoing reports, given academic research tends to move at a slower pace than industry timelines do.

Researchers, in trying to meet the needs of their tech worker participants, should also carefully outline their terms in regards to the research, else the balance may tip entirely in favor of the company. For example, I had conducted an interesting and novel study on the perceived importance of identity characteristics in computer vision for a company, but I did not have permission

---

[6]Note: this is an observable bias in many other human subjects studies [24]

to publish it and colleagues at the company had no incentive to collaborate on doing so.

6.2 Opportunity 3: How Industry can Provide Meaningful Incentives for Research Participation

Currently, participating in third-party academic research or even collaborating on second-party research is generally not incentivized for the majority of tech workers. Only those in research roles who collaborate on research projects with scholarly output are incentivized by their organizations in the form of KPIs. Marijan and Sen identify a number of solutions for the bridging "research–practice collaboration gap," including improving reciprocity between industry stakeholders and academic researchers [53]. One of their suggestions includes encouraging industry partners and researchers to work together to meet both research goals and KPIs. To better promote tech worker participation in research, companies could also consider ways of meaningfully rewarding participation:

- Many organizations include the concept of "impact" in their KPIs. Companies might consider research publications and reports which workers participated in as external impact. They might track such impact by creating forms to provide proof of participation, even when that research is anonymized. For those workers who are measured by participation in service activities, companies might consider expanding the concept of service activities to include research participation.
- Offering research collaboration trainings, which focus on teaching employees in a variety of roles how to work with researchers to benefit their outputs and increase their impacts.
- Some companies incentivize their workers to participate in certain activities by rewarding them small gifts for completing tasks. For example, some organizations have "wellness programs," and provide financial or other incentives for activities such as exercising [12, 60]. Companies might consider ways of compensating participants for research activities under an ethos of transparency and open science.

Of course, incentives for academic research extend beyond individual employees. Organizations themselves must see benefits for participating in and collaborating on research. Companies can view external research activities as methods for increasing their transparency, showcasing their commitment to knowledge sharing, and even gaining novel insights to help their business practices without having to directly pay for it (in the form of salaries or grants).

7 CHALLENGE 4: (SUB)CONTRACTING AND COMPLIANCE REVIEWS CAN BLOCKADE SECOND-PARTY (AND EVEN FIRST-PARTY) RESEARCH

Second-party research can be an effective way to overcome the limitations and opacity of third-party research. However, second-party research also comes with unique limitations which prevent access and publication. Even as an external researcher given internal access to a company, I encountered three barriers which were nearly insurmountable: (1) contracts that prevented contingent workers from being involved in research; (2) company policies which obscured the identities of clients; and (3) the blockading of research projects to protect the company from legal liability or scandal.

First, contingent workers who are directly employed by a tech company are, often, entirely inaccessible to researchers. Researchers might think that getting access to contingent workers, like data annotators, would be more accessible if they are an insider to a company. However, the contracts contingent workers sign prevent second-party and even first-party researchers from interacting with them. For one study, I had hoped to get access to data annotators employed directly by tech companies. I worked with colleagues internal to a large tech company, who contacted numerous data teams and lawyers employed within the company to understand the mechanisms for speaking with data annotators. In the end, colleagues and I were told that it was impossible to speak with data annotators contracted by the company because their contracts prevented contacting individual data workers or collecting information from them, including attributes like race and gender. Conducting research with contracted data annotators would involve renegotiating employment contracts. As a third-party researcher I had no means to identify data workers who contracted for specific tech companies, and as a second-party researcher, I was explicitly denied access to data workers, limiting the types of research that can be done on data annotation work entirely.

Second, the identities of clients were often formally (through the use of confidentiality agreements and subcontracting [36]) or informally kept secret. While some of the clients who purchase products and services from tech companies are relatively open about the existence of a relationship (e.g., through usually vague testimonials on company websites), tech workers largely declined to make connections with the customer representatives they have worked with, stating, for example, *"we need to protect our customers."* Beyond refusing to make connections for third- or second-party researchers, companies often do not even actively track how their customers are using their services (e.g., APIs) due to privacy concerns. This meant that first-party tech workers also largely did not know the use cases that their models were being deployed for. Not only does this make it difficult for researchers to triangulate around specific products or business practices, it makes knowing when, where, how, and by whom a company's products are being used murky, at best.

Finally, companies employ compliance review processes before and after second- and first-party researchers (as collaborators) can conduct or publish research, in any official capacity. These processes are often dubbed "legal review," "ethics review," and/or "comms review." Legal review is generally aimed at ensuring that research complies with the law and internal company policy; ethics review is generally aimed at ensuring the research follows ethical standards and values (much like institutional IRBs); and comms (communications) review is generally aimed at ensuring that research publications do not violate NDAs, reveal trade secrets, or paint the company in a bad light. The use of these different forms of reviews might differ from company to company, and many of the intended uses of them tend to overlap (e.g., ethics review might also be a form of comms review).

For second-party researchers like myself, I discovered how compliance reviews can prevent access to certain types of data on certain types of populations. Unlike IRBs, which do not wholesale ban the research of specific groups or the collections of certain types of data, company reviews often do. For example, I collaborated with an industry research team to conduct a study understanding how moderators react to the types of content they are employed to moderate. At this stage we were considering the viability of measuring the harmfulness of online content. However, the company which was providing access to moderators denied approval for this study design, due to concerns that knowing certain information might make the company vulnerable to legal liability. The company also worried that any internal leaks could be especially damaging, given that the wellbeing of moderators was a salient conversation in news outlets at the time (e.g., [57]).

7.1 Navigating Challenge 4: How Researchers can Reframe Research Goals When Blockaded by Companies

In Challenge 4, I described three mechanisms which bar certain types of research even when an external researcher has second-party access—research on certain populations (contingent workers, clients) and on certain topics (those deemed too risky by compliance review teams). In the context of my own experiences, **I had not identified ways to overcome the challenge directly**. Thus, I will describe how I navigated each of these three mechanisms—–largely by pivoting research directions.

7.1.1 Use proxies for inaccessible research populations. Researchers who have identified difficult populations to secure can also consider potential proxies for those specific populations as viable alternatives. Given there were no means for myself or my first-party industry contacts to access contingent data workers, I instead chose to seek access to data workers elsewhere. As previously mentioned, I instead worked to build a relationship with a data production company, who might provide services to companies like those my industry contacts worked with. I also chose to work with freelance data workers on Upwork, putting out a job ad for data annotators who completed jobs for private companies rather than academics. While these workers were unable to name their clients due to their own NDAs, they also had experience working on projects which might be generalizable to the tech industry at large. Further, having built a relationship with this data production company, I was then able to get access to some of their clients as a second-party researcher—a population I could not reach as a third-party researcher or as a second-party researcher in a large tech company. Even if researchers have to pivot from their initial population, they may build relationships that open different doors during the research process.

7.1.2 Be flexible with research design for the sake of maintaining second-party research access. Given the constraints that come up around both research populations and topics, second-party researchers should consider alternative methods for answering their research questions. For example, in attempting to understand how content moderators perceived the harmfulness of the content they

work with, myself and collaborators were denied access by compliance review before we could run the study. Thus, we instead chose to use more inductive qualitative methods which centered moderator opinions. We pivoted our methods away from conducting physiological laboratory studies to doing interviews, surveys, and card sorting activities. The study was then approved and we were given access to moderators to work with. If researchers must go through a company to conduct research they find important, they should be prepared to negotiate the exact studies they can do.

7.2 Opportunity 4: How Industry can Reconsider its Approach to Legal Liability and Public Relations in its (Sub)contracts and Compliance Reviews

Candidly, Challenge 4 was the most difficult to navigate, to the point where "navigating" meant pivoting to proxies for the participants I originally had in mind or outright shifting my research questions. Thus, Challenge 4 requires a lot more substantive change than previous challenges. Academic researchers must weigh the importance of research insights with what industry makes accessible. Industry acting as gatekeepers to specific stakeholders or topics means that, as it stands now, researchers have no choice but to either accept limited access or to pivot to another population or research question. The practice of using internal reviews to deny studies the broader community believes is important has also come under fire recently (e.g., [66]). Even when researchers are given second-party access to corporations, that research may reflect the corporate interests of the company [38]. Further, it is difficult for external reviewers to assess the trustworthiness of corporate research due to internal review processes and the economic incentives researchers might have to maintain relationships with industry [40, 90].

While tech is increasingly operating "data clean rooms," where anonymized user data can be shared with external researchers, there are no mechanisms for obtaining access to client or contractor data or participants. Regardless of the source or type of data, tech companies have options for balancing privacy and transparency. Young et al. present both legal and technical frameworks for preserving privacy while conducting transparent research [91]. Such approaches can be extended beyond user data to encompass data collected on workers and clients. Below are considerations for loosening the restrictive mechanisms governing research in industry:

- Considering the increased interest around the rights and wellbeing of contingent workers (e.g., [18, 47, 93]), companies should reconsider excluding research opportunities from their contracts. Knowledge about the role of demographics, worker perspectives, and labor conditions can be useful for companies to understand everything from potential sources of biases in their models to opportunities to improve policies around tasks like data labeling and content moderation.
- Companies should consider being more public about their client contracts, even if that is simply listing the businesses using a given product on their websites. Given the discovery of dubious connections between tech company products and their use cases (e.g., [92]), companies would benefit from knowing more about how their products are actively

being used. Researchers can also provide deeper insight into company-client relationships that extend beyond business deals, such as unstated latent needs [70].

- Review boards (ethics/legal/comms reviews) should focus their effort on anonymization of company identity and IP, rather than silencing specific research topics or potentially critical findings. They should instead view findings as opportunities to improve their practices. Companies can also consider when it is beneficial to the public and their brand to be public about research participation.

Unlike some of the opportunities focused on dealing with tech worker anxieties about their NDAs (see Opportunity 2), addressing Challenge 4 would require tech companies to actively reconsider their approach to risk management and relationships and contracts with clients and contractors. Given this is unlikely to change for the sake of academic research, policy and law once more become an avenue for increasing transparency. For example, it might be time to reconsider allowing tech companies to conduct their own internal ethics reviews, and instead expand the role of federal oversight of research in ways which prioritize the public, technology users, and workers [89].

8 CONCLUSION AND FUTURE WORK

External research on the companies developing AI is crucial to ensuring transparency, grounding interventions in informed real-world policies and procedures, and unearthing opaque perspectives and social norms that feed into the development context—ideally, without the types of biases that might permeate internally conducted studies. However, conducting empirical research on individual and organizational level practices, policies, and perspectives in the tech industry is difficult. Drawing from research exploring the development of algorithmic systems, I identified four challenges to conducting third- and second-party research on the tech industry. These four challenges spanned recruitment to publication stages. For each challenge, I described how I navigated it, ideally providing methodological guidance and advice to other researchers facing similar challenges. Finally, for each challenge I provide potential opportunities for the tech industry itself to intervene and ease restrictions for external researchers, ideally appealing to the priorities of private companies. Naturally, even if companies were to adopt such opportunities there would still be trade-offs. After all, openly letting academic researchers into the walled garden would reflect corporate interests to some degree, thus continuing already present concerns about trustworthiness and economic conflicts of interest.

Further, I also acknowledge that the priorities of the tech industry likely outweigh the importance of open and transparent research. Both formal and cultural norms within companies drive workers to focus most on creating marketable products, maintaining client relationships, and protecting the company from reputational and legal harm. Thus, I openly acknowledge that to increase the transparency of the tech industry, legal regulation is likely the most pivotal step. Currently, academia, though an industry in itself, is held to more rigorous research standards than private companies. Yet private companies, especially Big Tech, have more wide-reaching implications for human subjects—the general public, across the globe.

Therefore, in the ongoing search for how best to regulate the development of AI and its influence on public life, regulators should also consider the human elements of AI development, the workers driving that development from the ground up.

Beyond legal intervention, there are still vast opportunities for further understanding the challenges of conducting research on the tech industry. Studies with third-, second-, and even first-party researchers could uncover many more challenges, methods for navigating them, and opportunities for change. There is also opportunity to understand how industry stakeholders themselves view the merit (or lack thereof) of participating in external research and what solutions they see for improving industry-academic relationships. I plan to conduct future work specifically aimed at understanding experiences with compliance review processes, so I might ground better insights on how such review processes could change.

ACKNOWLEDGMENTS

REFERENCES

[1] Sanna J. Ali, Angèle Christin, Andrew Smart, and Riitta Katila. 2023. Walking the Walk of AI Ethics: Organizational Challenges and the Individualization of Risk among Ethics Entrepreneurs. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 217–226. https://doi.org/10.1145/3593013.3593990

[2] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks. *ProPublica.org* (2016), 1–17.

[3] Rachel S. Arnow-Richman, Gretchen Carlson, Orly Lobel, Julie Roginsky, Jodi L. Short, and Evan Starr. 2022. Supporting Market Accountability, Workplace Equity, and Fair Competition by Reining in Non-Disclosure Agreements.

[4] Chelsea Barabas, Colin Doyle, JB Rubinovitz, and Karthik Dinakar. 2020. Studying up: Reorienting the Study of Algorithmic Fairness around Issues of Power. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 167–176. https://doi.org/10.1145/3351095.3372859

[5] Teanna Barrett, Quanze Chen, and Amy Zhang. 2023. Skin Deep: Investigating Subjectivity in Skin Tone Annotations for Computer Vision Benchmark Datasets. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 1757–1771. https://doi.org/10.1145/3593013.3594114

[6] Abeba Birhane and Vinay Uday Prabhu. 2021. Large Image Datasets: A Pyrrhic Win for Computer Vision?. In *Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision, WACV 2021*. Institute of Electrical and Electronics Engineers Inc., 1536–1546. https://doi.org/10.1109/WACV48630.2021.00158 arXiv:2006.16923

[7] Emily Birnbaum. 2020. A Wall of Silence Holding Back Racial Progress in Tech: NDAs. *Protocol* (July 2020).

[8] Hannah Bloch-Wehba. 2023. The Promise and Perils of Tech Whistleblowing.

[9] William Boag, Harini Suresh, Bianca Lepe, and Catherine D'Ignazio. 2022. Tech Worker Organizing for Power and Accountability. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 452–463. https://doi.org/10.1145/3531146.3533111

[10] Shannon Bond. 2021. NYU Researchers Were Studying Disinformation On Facebook. The Company Cut Them Off. *NPR* (Aug. 2021).

[11] Joy Buolamwini and Timnit Gebru. 2018. *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification **. Technical Report. 1–15 pages.

[12] CDC. 2023. What's Your Role? Employers. https://www.cdc.gov/physicalactivity/activepeoplehealthynation/everyone-can-be-involved/employers.html.

[13] Kathy Charmaz. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*.

[14] Elizabeth Chase. 2017. Enhanced Member Checks: Reflections and Insights from a Participant-Researcher Collaboration. *The Qualitative Report* 22, 10 (Oct. 2017), 2689–2703. https://doi.org/10.46743/2160-3715/2017.2957

[15] Shanley Corvite, Kat Roemmich, Tillie Ilana Rosenberg, and Nazanin Andalibi. 2023. Data Subjects' Perspectives on Emotion Artificial Intelligence Use in the Workplace: A Relational Ethics Lens. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (April 2023), 124:1–124:38. https://doi.org/10.1145/3579600

[16] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1571–1583. https://doi.org/10.1145/3531146.3533213

[17] Paula Czarnowska, Yogarshi Vyas, and Kashif Shah. 2021. Quantifying Social Biases in NLP: A Generalization and Empirical Comparison of Extrinsic Fairness Metrics. *Transactions of the Association for Computational Linguistics* 9 (Nov. 2021), 1249–1267. https://doi.org/10.1162/tacl_a_00425

[18] Kaitlin Daniels and Michal Grinstein-Weiss. 2018. The Impact of the Gig-Economy on Financial Hardship Among Low-Income Families. *SSRN Electronic Journal* (June 2018). https://doi.org/10.2139/ssrn.3293988

[19] Claes de Vreese and Rebekah Tromble. 2023. The Data Abyss: How Lack of Data Access Leaves Research and Society in the Dark. *Political Communication* 40, 3 (May 2023), 356–360. https://doi.org/10.1080/10584609.2023.2207488

[20] Wesley Hanwen Deng, Nur Yildirim, Monica Chang, Motahhare Eslami, Kenneth Holstein, and Michael Madaio. 2023. Investigating Practices and Opportunities for Cross-functional Collaboration around AI Fairness in Industry Practice. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 705–716. https://doi.org/10.1145/3593013.3594037

[21] Emily Denton, Alex Hanna, Razvan Amironesei, Andrew Smart, and Hilary Nicole. 2021. On the Genealogy of Machine Learning Datasets: A Critical History of ImageNet. *Big Data and Society* 8, 2 (Sept. 2021), 205395172110359. https://doi.org/10.1177/20539517211035955

[22] Desjardins. 2022. Congress Passes Law Banning Non-Disclosure Agreements in Sexual Harassment Cases. *PBS NewsHour* (Nov. 2022).

[23] Matt Drange. 2021. Apple Agrees to "Make Improvements" to Its NDAs after Whistleblower Documents Showed Evidence the Company Used the Agreements to Silence Employees. *Business Insider* (Nov. 2021).

[24] Dirk M. Elston. 2021. Participation Bias, Self-Selection Bias, and Response Bias. *Journal of the American Academy of Dermatology* 0, 0 (June 2021). https://doi.org/10.1016/j.jaad.2021.06.025

[25] Michael Feffer, Nikolas Martelaro, and Hoda Heidari. 2023. The AI Incident Database as an Educational Tool to Raise Awareness of AI Harms: A Classroom Exploration of Efficacy, Limitations, & Future Improvements. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '23)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3617694.3623223

[26] Virginia Felkner, Ho-Chun Herbert Chang, Eugene Jang, and Jonathan May. 2023. WinoQueer: A Community-in-the-Loop Benchmark for Anti-LGBTQ+ Bias in Large Language Models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (Eds.). Association for Computational Linguistics, Toronto, Canada, 9126–9140. https://doi.org/10.18653/v1/2023.acl-long.507

[27] Deen Freelon. 2018. Computational Research in the Post-API Age. *Political Communication* 35, 4 (Oct. 2018), 665–668. https://doi.org/10.1080/10584609.2018.1477506

[28] Vinitha Gadiraju, Shaun Kane, Sunipa Dev, Alex Taylor, Ding Wang, Emily Denton, and Robin Brewer. 2023. "I Wouldn't Say Offensive but...": Disability-Centered Perspectives on Large Language Models. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 205–216. https://doi.org/10.1145/3593013.3593989

[29] Hadass Goldblatt, Orit Karnieli-Miller, and Melanie Neumann. 2011. Sharing Qualitative Research Findings with Participants: Study Experiences of Methodological and Ethical Dilemmas. *Patient Education and Counseling* 82, 3 (March 2011), 389–395. https://doi.org/10.1016/j.pec.2010.12.016

[30] GPAI. 2023. *Fairwork AI Ratings 2023: The Workers Behind AI at Sama*. Technical Report. Global Partnership on AI, Oxford: United Kingdom.

[31] Mary L. Gray and Suri Siddharth. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*.

[32] Lara Groves. 2023. *Going Public: Exploring Public Participation in Commercial AI Labs*. Technical Report. Ada Lovelace Institute.

[33] Luke Guerdan, Amanda Coston, Zhiwei Steven Wu, and Kenneth Holstein. 2023. Ground(Less) Truth: A Causal Framework for Proxy Labels in Human-Algorithm Decision-Making. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 688–704. https://doi.org/10.1145/3593013.3594036

[34] Ronald E. Hallett. 2013. Dangers of Member Checking. In *The Role of Participants in Education Research: Ethics, Epistemologies, and Methods*. Routledge.

[35] Foad Hamidi, Morgan Klaus Scheuerman, and Stacy M Branham. 2018. Gender Recognition or Gender Reductionism? The Social Implications of Automatic Gender Recognition Systems. In *2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*.

[36] Glen Harvey. 2020. Thousands of Contracts Highlight Quiet Ties between Big Tech and U.S. Military. *NBC News* (July 2020).

[37] Kenneth Holstein, Hal Daumé III, Miroslav Dudík, Hanna Wallach, and Jennifer Wortman Vaughan. 2019. Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 16. https://doi.org/10.1145/3290605.3300830 arXiv:1812.05239v1

[38] Hubert Horan. 2019. Uber's "Academic Research" Program: How to Use Famous Economists to Spread Corporate Narratives.

[39] Camilla Alexandra Hrdy and Christopher B. Seaman. 2023. Beyond Trade Secrecy: Confidentiality Agreements That Act Like Noncompetes. https://doi.org/10.2139/ssrn.4384661

[40] Lilly Irani, Niloufar Salehi, Joyojeet Pal, Andrés Monroy-Hernández, Elizabeth Churchill, and Sneha Narayan. 2019. Patron or Poison? Industry Funding of HCI Research. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*. Association for Computing Machinery, 111–115. https://doi.org/10.1145/3311957.3358610

[41] Abigail Z. Jacobs and Hanna Wallach. 2019. *Measurement and Fairness*. Technical Report. arXiv:1912.05511

[42] Khari Johnson. 2020. Google Cloud AI Removes Gender Labels from Cloud Vision API to Avoid Bias. *VentureBeat* (Feb. 2020).

[43] Stephen Joyce. 2022. States Act to Curb Employer Overuse of Non-Disclosure Agreements. *Bloomberg Law* (Sept. 2022).

[44] Pratyusha Ria Kalluri, William Agnew, Myra Cheng, Kentrell Owens, Luca Soldaini, and Abeba Birhane. 2023. The Surveillance AI Pipeline.

[45] Yarden Katz. 2020. *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*. Columbia University Press.

[46] Jared Katzman, Angelina Wang, Morgan Klaus Scheuerman, Su Lin Blodgett, Kristen Laird, Hanna Wallach, and Solon Barocas. 2023. Taxonomizing and Measuring Representational Harms: A Look at Image Tagging. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*.

[47] Boitumelo Lesala Khethisa, Pitso Tsibolane, and Jean Paul Van Belle. 2020. Surviving the Gig Economy in the Global South: How Cape Town Domestic Workers Cope. *IFIP Advances in Information and Communication Technology* 601 (2020), 67–85. https://doi.org/10.1007/978-3-030-64697-4_7/TABLES/3

[48] Gabriel Lima, Nina Grgic-Hlaca, Jin Keun Jeong, and Meeyoung Cha. 2023. Who Should Pay When Machines Cause Harm? Laypeople's Expectations of Legal Damages for Machine-Caused Harm. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 236–246. https://doi.org/10.1145/3593013.3593992

[49] Orly Lobel. 2018. NDAs Are Out of Control. Here's What Needs to Change. *Harvard Business Review* (Jan. 2018).

[50] Ryan Mac and Sheera Frenkel. 2021. Facebook Downplays Internal Research Released on Eve of Hearing. *The New York Times* (Sept. 2021).

[51] Matt MacNeil. 2021. Ziad Obermeyer and Colleagues at the Booth School of Business Release Health Care Algorithmic Bias Playbook. https://publichealth.berkeley.edu/news-media/research-highlights/ziad-obermeyer-and-colleagues-at-the-booth-school-of-business-release-health-care-algorithmic-bias-playbook/.

[52] Vidushi Marda and Shivangi Narayan. 2020. Data in New Delhi's Predictive Policing System. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 317–324. https://doi.org/10.1145/3351095.3372865

[53] Dusica Marijan and Sagar Sen. 2022. Industry–Academia Research Collaboration and Knowledge Co-creation: Patterns and Anti-patterns. *ACM Transactions on Software Engineering and Methodology* 31, 3 (March 2022), 45:1–45:52. https://doi.org/10.1145/3494519

[54] Milagros Miceli, Julian Posada, and Tianling Yang. 2022. Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power?. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 6. ACM PUB27 New York, NY, USA. https://doi.org/10.1145/3492853

[55] Laura Nader. 1972. *Up the Anthropologist: Perspectives Gained From Studying Up*. Technical Report.

[56] Kelly A. Negrin, Susan E. Slaughter, Sherry Dahlke, and Joanne Olson. 2022. Successful Recruitment to Qualitative Research: A Critical Reflection. *International Journal of Qualitative Methods* 21 (April 2022), 16094069221119576. https://doi.org/10.1177/16094069221119576

[57] Casey Newton. 2019. The Secret Lives of Facebook Moderators in America. *The Verge* (2019).

[58] Ziad Obermeyer and Sendhil Mullainathan. 2019. Dissecting Racial Bias in an Algorithm That Guides Health Decisions for 70 Million People. In *Proceedings*

of the Conference on Fairness, Accountability, and Transparency - FAT '19*. ACM Press, New York, New York, USA, 89–89. https://doi.org/10.1145/3287560.3287593

[59] Rodrigo Ochigame. 2019. The Invention of "Ethical AI:" How Big Tech Manipulates Academia to Avoid Regulation. *The Intercept* (Dec. 2019).

[60] Renée Onque. 2023. This Company Pays Employees to Exercise: 'If They Come to Those Workouts, They Are on the Clock,' Says CEO. *CNBC* (May 2023).

[61] Victoria Pagan. 2021. The Murder of Knowledge and the Ghosts That Remain: Non-Disclosure Agreements and Their Effects. *Culture and Organization* 27, 4 (July 2021), 302–317. https://doi.org/10.1080/14759551.2021.1907389

[62] Carol Passos, Daniela S. Cruzes, Tore Dybå, and Manoel Mendonça. 2012. Challenges of Applying Ethnography to Study Software Practices. In *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '12)*. Association for Computing Machinery, New York, NY, USA, 9–18. https://doi.org/10.1145/2372251.2372255

[63] Deborah Raji, Emily Denton, Emily M. Bender, Alex Hanna, and Amandalynne Paullada. 2021. AI and the Everything in the Whole Wide World Benchmark. *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks* 1 (Dec. 2021).

[64] Inioluwa Deborah Raji and Joy Buolamwini. 2019. *Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products*. Technical Report. 7 pages.

[65] Bogdana Rakova, Jingying Yang, Henriette Cramer, and Rumman Chowdhury. 2021. Where Responsible AI Meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1 (April 2021). https://doi.org/10.1145/3449081

[66] Betsy Reed. 2021. Google to Change Research Process after Uproar over Scientists' Firing. *The Guardian* (Feb. 2021).

[67] Cecilia Rikap and Bengt-Åke Lundvall. 2022. Big Tech, Knowledge Predation and the Implications for Development. *Innovation and Development* 12, 3 (Sept. 2022), 389–416. https://doi.org/10.1080/2157930X.2020.1855825

[68] Mark Ryan, Eleni Christodoulou, Josephina Antoniou, and Kalypso Iordanou. 2022. An AI Ethics 'David and Goliath': Value Conflicts between Large Tech Companies and Their Employees. *AI & SOCIETY: Journal of Knowledge, Culture and Communication* (March 2022), 1–16. https://doi.org/10.1007/s00146-022-01430-1

[69] Janet Salmons. 2018. *Doing Qualitative Research Online*. SAGE Publications Ltd. https://doi.org/10.4135/9781473921955

[70] Edmund Christian Salzmann and Alexander Kock. 2020. When Customer Ethnography Is Good for You – A Contingency Perspective. *Industrial Marketing Management* 88 (July 2020), 366–377. https://doi.org/10.1016/j.indmarman.2020.05.027

[71] Morgan Klaus Scheuerman. 2022. Envisioning Identity: The Social Production of Human-Centric Computer Vision Systems. In *Companion Publication of the 2022 Conference on Computer Supported Cooperative Work and Social Computing*. Association for Computing Machinery, New York, NY, USA, 210–213. https://doi.org/10.1145/3500868.3561396

[72] Morgan Klaus Scheuerman and Jed R. Brubaker. 2024-05-11/2024-05-16. Products of Positionality: How Tech Workers Shape Identity Concepts in Computer Vision. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu, HI. https://doi.org/10.1145/3613904.3641890

[73] Morgan Klaus Scheuerman, Jialun Aaron Jiang, Casey Fiesler, and Jed R Brubaker. 2021. A Framework of Severity for Harmful Content Online. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021). https://doi.org/10.1145/3479512

[74] Morgan Klaus Scheuerman, Madeleine Pape, and Alex Hanna. 2021. Auto-Essentialization: Gender in Automated Facial Analysis as Extended Colonial Project. *Big Data & Society* 8, 2 (Dec. 2021), 205395172110537. https://doi.org/10.1177/20539517211053712

[75] Morgan Klaus Scheuerman, Jacob M Paul, and Jed R Brubaker. 2019. How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis and Image Labeling Services. In *Proc. ACM Hum.-Comput. Interact.*, Vol. 144. Association for Computing Machinery, 33. https://doi.org/10.1145/3359246

[76] Olivia Solon. 2020. Facebook Management Ignored Internal Research Showing Racial Bias, Current and Former Employees Say. *NBC News* (July 2020).

[77] Howie Stein, Goutham Belliappa, and Ben Coffey. 2023. Data Clean Rooms Could Create New Opportunities for Marketers. *WSJ* (Jan. 2023).

[78] Paola Tubaro. 2021. Whose Results Are These Anyway? Reciprocity and the Ethics of "Giving Back" after Social Network Research. *Social Networks* 67 (Oct. 2021), 65–73. https://doi.org/10.1016/j.socnet.2019.10.003

[79] Chiara Ullstein, Severin Engelmann, Orestis Papakyriakopoulos, Michel Hohendanner, and Jens Grossklags. 2022. AI-Competent Individuals and Laypeople Tend to Oppose Facial Analysis AI. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '22)*. Association for Computing Machinery, New York, NY, USA, Article 9. https://doi.org/10.1145/3551624.3555294

[80] Kees van der Waal. 2009. Organizational Ethnography: Studying the Complexities of Everyday Life. In *Organizational Ethnography: Studying the Complexities of Everyday Life*. SAGE Publications Ltd, London, 23–39. https://doi.org/10.4135/9781446278925

[81] Niels van Doorn. 2017. Platform Labor: On the Gendered and Racialized Exploitation of Low-Income Service Work in the 'on-Demand' Economy. *Information Communication and Society* 20, 6 (June 2017), 898–914. https://doi.org/10.1080/1369118X.2017.1294194

[82] Michael Veale, Max Van Kleek, and Reuben Binns. 2018. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (2018). https://doi.org/10.1145/3173574

[83] James Vincent. 2019. The Problem with AI Ethics. *The Verge* (April 2019).

[84] David Gray Widder, Sarah West, and Meredith Whittaker. 2023. Open (For Business): Big Tech, Concentrated Power, and the Political Economy of Open AI. https://doi.org/10.2139/ssrn.4543807

[85] David Gray Widder, Derrick Zhen, Laura Dabbish, and James Herbsleb. 2023. It's about Power: What Ethical Concerns Do Software Engineers Have, and What Do They (Feel They Can) Do about Them?. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. Association for Computing Machinery.

[86] Adrienne Williams, Milagros Miceli, and Timnit Gebru. 2022. The Exploited Labor behind Artificial Intelligence. *NOEMA* (2022), 1–11.

[87] Sabrina Willmer, Austin Weinstein, and Bloomberg. 2023. 'It's Concerning': Tech Firms Are Using NDAs to Illegally Muzzle Whistleblowers by Threatening to Sue Them for Talking, SEC Says. *Fortune* (April 2023).

[88] Richmond Y. Wong. 2021. Tactics of Soft Resistance in User Experience Professionals' Values Work. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 355:1–355:28. https://doi.org/10.1145/3479499

[89] Karen Yeung, Andrew Howes, and Ganna Pogrebna. 2020. AI Governance by Human Rights–Centered Design, Deliberation, and Oversight: An End to Ethics Washing. In *The Oxford Handbook of Ethics of AI*, Markus D. Dubber, Frank Pasquale, and Sunit Das (Eds.). Oxford University Press, 0. https://doi.org/10.1093/oxfordhb/9780190067397.013.5

[90] 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

[91] Meg Young, Luke Rodriguez, Emily Keller, Feiyang Sun, Boyang Sa, Jan Whittington, and Bill Howe. 2019. Beyond Open vs. Closed: Balancing Individual Privacy and Public Accountability in Data Sharing. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 191–200. https://doi.org/10.1145/3287560.3287577

[92] Zak Doffman. 2019. Is Microsoft AI Helping To Deliver China's 'Shameful' Xinjiang Surveillance State? *Forbes* (2019).

[93] Zsolt Ződi and Bernát Török. 2021. Constitutional Values in the Gig-Economy? Why Labor Law Fails at Platform Work, and What Can We Do about It? *Societies* 11, 3 (2021). https://doi.org/10.3390/soc11030086