Building, Shifting, & Employing Power: A Taxonomy of Responses From Below to Algorithmic Harm

Alicia DeVrio
Carnegie Mellon University
Pittsburgh, PA, USA
adevos@andrew.cmu.edu

Motahhare Eslami*
Carnegie Mellon University
Pittsburgh, PA, USA
meslami@andrew.cmu.edu

Kenneth Holstein*
Carnegie Mellon University
Pittsburgh, PA, USA
kjholste@andrew.cmu.edu

ABSTRACT

A large body of research has attempted to ensure that algorithmic systems adhere to notions of fairness and transparency. Increasingly, researchers have highlighted that mitigating algorithmic harms requires explicitly taking power structures into account. Those with power over algorithmic systems often fail to sufficiently address algorithmic harms and rarely consult those directly harmed by algorithmic systems. Left to their own devices, people respond to algorithmic harms they encounter in a wide variety of ways, but we lack broader, overarching understandings of these responses. In this work, we synthesize documented, historical cases into a taxonomy of responses "from below" to algorithmic harm. Our taxonomy connects different types of responses to existing theorizations of power from fields including anthropology, human-computer interaction, and communication, centering how people employ, shift, and build power in their responses to algorithmic harm. Based on our taxonomy, we highlight an opportunity space for the FAccT community to engage with and support such action from below.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; **Collaborative and social computing theory, concepts and paradigms**.

KEYWORDS

algorithmic harm, power, algorithmic bias, critical algorithm studies, data leverage, feminist refusal, algorithmic resistance, everyday algorithm auditing, AI ethics, accountability

ACM Reference Format:
Alicia DeVrio, Motahhare Eslami, and Kenneth Holstein. 2024. Building, Shifting, & Employing Power: A Taxonomy of Responses From Below to Algorithmic Harm. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24), June 03–06, 2024, Rio de Janeiro, Brazil.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3630106.3658958

*Co-senior authors contributed equally to this research.

FAccT '24, June 03–06, 2024, Rio de Janeiro, Brazil

ACM ISBN 979-8-4007-0450-5/24/06
https://doi.org/10.1145/3630106.3658958

1 INTRODUCTION

Aiming to root out algorithmic harms, a large body of research has attempted to ensure that algorithmic systems adhere to notions of fairness and transparency. Increasingly though, researchers have highlighted that effectively addressing harmful algorithmic behaviors requires centering systems of power in and around algorithmic systems [17, 24, 25, 77, 82, 95]. This focus is especially important since existing structures around algorithmic systems often fail to empower those who directly interact with and are affected by algorithmic harm [17].

In response to such power imbalance, people affected by algorithmic harm frequently adopt workarounds and counter-practices to algorithmic systems in place. Recent research and real-world case studies indicate that responses "from below"—i.e., from members of the public or others without direct power to make decisions about the design or deployment of algorithmic systems [142]—can be effective in garnering power and working toward harm remediation [45, 85, 120, 137, 142]. For example, Twitter users who noticed, tweeted about, and extensively worked together to test racial bias in the platform's image cropping algorithm succeeded in causing Twitter to reconsider and change its use of the algorithm [120]. And after Uber and Lyft algorithmically determined inadequate ride fares, drivers coordinated turning their apps off to feign a driver shortage and were able to boost fares [85, 130].

While responses from below can be effective in mitigating some algorithmic harms, actors' structural disempowerment can limit the effectiveness of their responses. For example, Elon Musk disbanded Twitter's AI ethics team after taking over the company [41], and rideshare companies continue to perpetuate algorithmic wage discrimination [47]. How can we better understand the ways that responses from below interact with power? What lessons can we draw from past responses to empower future action from below?

In this paper, we develop a taxonomy of existing types of responses to algorithmic harms, centering actions that people take from below in response to algorithmic harms. We focus on the nature of the responses themselves and, where described by actors, their motivations and aims; we do not in this work attempt to categorize the effectiveness of various responses. Our taxonomy emphasizes how these responses from below are linked to power dynamics through three broad categories: "*protecting*" themselves by employing their own power, typically within the context of a harmful algorithmic system; "*pressuring*" those in control of algorithmic systems by shifting power away from them; and "*strengthening*" their capacity by building their reserves of power, enhancing their abilities to pressure and protect. Protecting responses can lead to mitigations more immediately, but they tend to occur within the existing logics of algorithmic systems, which inherently limits their

outcomes. For example, when users found Booking.com's hotel review algorithm problematic and acted to fix its outputs by manipulating their review scores, they were constrained by fundamental issues within the algorithm: its lowest possible output rating was 2.5, so altering inputs did not help when users wanted ratings lower than 2.5 [51]. More widespread, systemic change often comes from pressuring responses. For example, following worldwide protests of unfair algorithmically generated A-level exam scores, the algorithmic scores were revoked and replaced with teacher-determined ones [101]. And people need sufficient power, which strengthening responses build, to do any of this. For example, to escape algorithmic surveillance and control, food delivery workers in China created virtual communities that allowed them to share tactics to better navigate food delivery platforms [129].

Based on our taxonomy of responses, we discuss opportunities for the FAccT community to better engage with everyday peoples' current actions in response to algorithmic harm to further amplify actions they are already taking in productive ways. Our work identifies ways that structurally disempowered people create and further their own counterpractices intended to secure power over algorithmic systems. Supporting efforts from below can put those most affected and with little systemic power in control of remediation efforts, which can enable progress even when powerful actors like companies fail to see worth or incentivization in working toward algorithmic harm remediation. In addition to efforts to mitigate harms from above (e.g., by supporting developers in assessing and addressing issues), it is critical that the FAccT community helps affected people defend themselves from actively harmful systems.

2 RELATED WORK

2.1 Harm in Algorithmic Systems

Algorithmic harm is "the adverse lived experiences resulting from [an algorithmic] system's deployment and operation in the world" [119]. With the immense amount of harm produced by algorithmic systems has come considerable work aiming to understand these harms, teasing out the distinctions around types of algorithmic harm in order to direct algorithmic harm mitigation efforts in more fruitful directions (e.g., [19, 25, 34, 111, 119, 141]). For instance, one recent taxonomy outlines five high-level categories of algorithmic harm—representational, allocative, quality-of-service, interpersonal, and social system harms—with 20 subcategories [119].

Many efforts—in academia, industry, governmental sectors, and beyond—have attempted to ensure that algorithmic systems adhere to notions of fairness in order to root out algorithmic harms such as bias (e.g., [3, 39]). For example, a well-known distinction between allocative and representational harms arose from research on algorithmic bias [20, 40].

However, researchers have highlighted that we cannot productively address harmful algorithmic behaviors like bias without centering systems of power [17, 24, 25, 77, 82, 95]. For example, Miceli et al. argue that we have not yet sufficiently centered issues of power and explain how seeing the world through only the lenses of bias and fairness neglects power [95]. Alkhatib similarly contends that algorithmic harm derives from structural power and thus harm reduction requires withdrawing power from algorithmic systems [17]. In this work, we focus on power as a way to get at issues of algorithmic harm.

2.2 Power Dynamics & Addressing Algorithmic Harm

Throughout this paper we understand power as a multifaceted concept that can emanate from multiple sources [84]. HCI researchers have highlighted the importance of social and political theory conceptualizations of power as "power-over others" and as "power-to do" [18, 112]. We adopt these framings, focusing on the ways that actors from below retain, maintain, and create both (1) ways to get others to engage in actions that diminish algorithmic harms that they encounter and (2) ways to act directly to shape their experiences with algorithmic systems. Following past work on power within and around algorithmic systems [77, 95], we focus on the structurally influenced relationships between different groups of people [18]. Our power-aware view enables deeper understanding of how actors from below, responding to algorithmic harm, seek to interact with, garner, and use power.

Organizations in charge of algorithmic systems often fail to address algorithmic harms sufficiently of their own impetus [7, 67, 103, 131]. Frequently companies respond to concerns surrounding algorithmic harm only when faced with significant evidence and negative repercussions—and even then, often only provide empty statements. For example, families who criticized Amazon's auto-recommendations of materials used for suicide were acknowledged with an Amazon statement that "it made customer safety a top concern" [67]. However, Amazon continued to sell the relevant products, claiming they were not responsible for customers' misuse [131]. Clearly those affected cannot always count on organizations to remediate harms they introduce with their algorithmic systems. Structural power often fails to empower those who directly interact with and are affected by algorithmic systems, rarely allowing them input into these systems [17]. Even when everyday people are enabled to participate in AI, rarely are they able to control decisions themselves [38, 125].

In the face of this, collective action presents a way for people to work toward shared goals by banding together. A significant body of existing research examines the power of collective action to bring about positive change and investigates the ways that researchers can support collective action (e.g., [23, 87, 92, 106, 136, 143]). However, there are many unknowns around the best ways to employ collective action within the context of algorithmic systems.

As such, some emerging research has investigated how to empower affected people in their efforts at algorithmic harm remediation. For example, Protective Optimization Technologies (POTs) provide ways for affected parties to "correct, shift, or expose harms" in algorithmic systems without the cooperation of service providers [85]. Data leverage, which presents ways for people to wield their own input data to harm or support technologies and associated companies [137], can occur in the context of algorithmic systems. We orient our work accordingly and consider it in line with the notion of tactics from below in which people "build power over technology when power has not been given" [142].

2.3 Responses From Below to Algorithmic Harm

Significant research investigates how users and others affected by algorithmic systems perceive algorithmic harm (e.g., [45, 59, 90]). For example, DeVos et al. sought to understand how users identify, make sense of, and evaluate potentially harmful algorithmic behaviors [45]. In this work, we focus on steps that everyday people take after they have perceived harm in an algorithmic system.

Recent research has revealed that people with little structural power have unique abilities to respond to algorithmic harms [45, 85, 120, 137, 142]. With the prevalence of algorithmic harm, many of these people's responses can be considered forms of "everyday resistance", a continuous struggle implemented with "the ordinary weapons of relatively powerless groups" [113]. We also draw on "routine infrastructuring", as actors work to re-appropriate and build resilience around algorithmic systems [114]. For example, the concept of "everyday algorithm auditing" highlights how users, in their everyday interactions with algorithmic systems, can surface, interrogate, and *work toward remediation* of harmful algorithmic behaviors that more formal auditing methods fail to find [120].

However, we lack broader, overarching understandings of the nature of the existing responses from below to algorithmic harm. Some work has investigated the ways that users and others affected respond to harmful algorithmic behaviors (e.g., [51, 85, 123, 150]). Throughout, many questions have arisen regarding the nature of these remediative responses from below. In this work, we start to provide this insight, extending past literature around responses from below such as data leverage and POTs [85, 137]. However, instead of centering data or technologies, we center what the disempowered do now to respond to algorithmic harms and how their actions are rooted in dynamics of power.

3 POSITIONALITY STATEMENT

In this work, we attempt to understand existing efforts by disempowered people to respond to algorithmic harms they encounter. As academics, we have seen and experienced the ways in which power shapes our discourse, values, and priorities; as such, we focus here on the ways in which responses can elucidate how power dynamics are experienced and navigated by their actors. As junior scholars, we wield comparatively less power within academia while holding relatively high power as researchers to influence the ways decisions around algorithmic systems are made and considered. We understand our position as imbued with responsibility to use our work to shift power to others who are directly affected by algorithmic systems—a group that includes us, as we too have experienced some of these systems' harms firsthand—yet have little control over these systems. We aim to do this in this paper by bringing together and strengthening existing currents of thought and action around empowering everyday people. That being said, our aspirations to represent complexity in responses from below is limited by our set of experiences: we are a group of non-white, multiracial, and immigrant researchers who are nevertheless deeply conditioned by the implicit and explicit normalization of Western culture and hegemony where we currently live and work, even as we resist it. As such, we look forward to the ways in which others might expand, critique, reshape, and build on this work in ways that shift power even more to those who currently respond from below.

4 TAXONOMY OF RESPONSES FROM BELOW TO ALGORITHMIC HARM

In this section, we describe nine types of responses from below to algorithmic harm, organized into five broad categories (see Table 1). Our taxonomy centers *people* and captures the actions they take from below to *employ, shift, or build power* in response to algorithmic harms. Each subsection presents a *type of response* and examines power-related *strategies* associated with it, using a range of existing examples to illustrate the scope of each.

4.1 Process of Response Collection & Categorization

Our research team gathered 96 documented cases with 169 embedded responses from below that intend to remediate perceived algorithmic harm to create our taxonomy. The collection process started with an initial set, and expanded via iterative search of various resources including research literature, news media, and social media. Our cases spanned many different algorithmic domains (e.g., recommendation, image cropping, facial recognition) and types of harm (e.g., financial, stereotyping, surveillance). More collection details can be found in Appendix A.

We understand actors and actions from below as those lacking direct power to control the design and deployment of algorithmic systems. To emphasize power relations, we broadly refer to those who have control over algorithmic systems as *"algorithmlords"*, inspired by past work describing "data barons" who control data and "digital feudalism" in which tech companies behave as modern-day feudal lords [46, 62, 76]. In the vein of [142], we understand algorithmlords as those holding greater power to envision and construct algorithmic systems, occupying positions "above" those subjugated by their decisions, while acknowledging that relationships between people are more complex and nuanced than this fully captures (see [50, 69]).

We extracted and qualitatively coded responses from the cases we gathered, then conducted a bottom-up thematic analysis with a series of interpretation sessions on these responses. Based on this, we grouped our responses into *types* of responses based on similar responder behaviors present and, based on prior literature on power, synthesized these into three high-level *strategies*, each of which captures a different way that these responses aim to interact with power:

Protecting strategies: ways that people responding to algorithmic harm act from below to directly affect their or others' "power-to" do [112]. Typically this occurs within the context of the algorithmic platform itself and serves to make people's interactions with algorithmic systems safer. These responses fit into repair politics [71, 74]—that is, responses that work "through improvisations, patches and ingenuity, together with and within algorithmic systems" [132]—and more immediately improve the experiences and abilities of the disempowered.

Pressuring strategies: ways that people responding to algorithmic harm act from below in attempts to gain "power-over" algorithmlords [112]. These responses often have the ultimate aim

Table 1: Types of responses (bold), their strategies that interact with power (X), and examples

Response Type ↓	Protecting	Pressuring	Strengthening	Examples
Mitigating individual algorithmic harms	X			Changing search terms after problematic results [45], Engaging with hidden content to amplify it [79]
Pursuing legal avenues against algorithmic harm		X		Instituting policies to reclaim AI governance [10, 55], Litigating companies who deploy harmful algorithms [52]
Investigating potentially harmful algorithmic behaviors further			X	Researching background information [142], Testing potential issues [120]
Refusing legitimate engagement with harmful algorithmic systems				
Impairing an algorithmic system	X	X		Slashing tires of self-driving cars [105, 107], Unplugging privacy-invasive cameras [64, 148]
Depriving an algorithmic system	X	X		Logging out of Facebook [140], Leaving data-labelling job [98]
Confusing an algorithmic system	X	X		Mirroring, cropping, and filtering duplicate videos [145], Downloading software that automatically clicks ads [72]
Communicating algorithmic harm with others				
"Complaining" to those with more power	X	X		Contesting errors using a platform's appeals system [83], Submitting complaint to government agency [37, 42]
Publicizing algorithmic harms broadly		X	X	Tweeting about an issue [6, 13, 54, 120], Holding a press conference [142]
Sharing, discussing, and organizing with others affected by the algorithmic system	X		X	Describing in writing where others can see [51, 120], Creating, joining, and discussing in WeChat groups [147]

of persuading algorithmlords to make changes. However, they may vary in their degree of forcefulness, from punishing an organization with negative PR, to convincing an organization by flagging issues for them to address, to forcing a company with new legal requirements.

Strengthening strategies: ways that people build up their reserves of power, enhancing their future abilities to protect or pressure. Strengthening also encompasses ways that actors recognize and become better equipped to leverage their existing, often collective, power. Strengthening can help people understand the scope of the harm they have encountered, determine what should be done next, and gain perceived legitimacy.

We see the cases used as the basis for our taxonomy as illustrative but not necessarily exhaustive. Responses and types of responses commonly span multiple strategies. Additionally, due to the constantly changing and emerging nature of algorithmic systems, of their applications, of related harms, and of related responses to these harms, we see this work as an initial exploration and invite future researchers to expand on it. More process details can be found in Appendix A.

4.2 Mitigating Individual Algorithmic Harms

In this type of response, people adjust their own behaviors in and around algorithmic systems to change how the system works in a particular circumstance, without fundamentally changing how the algorithmic system works as a whole. We understand such actions as *protecting* responses, as people employ their own abilities to shield themselves from the harms a given algorithmic system foists upon them. These responses parallel conceptions of technological repair, which has been theorized as "the ongoing work of fixing and maintaining the objects and systems around us" [71] and has been previously applied to users addressing harmful algorithmic behaviors [51]. Additionally, these responses can be seen as a form of algorithmic resistance, theorized by Velkova and Kaun as action that "takes place from *within* the logic of the algorithm" and "does not deny the power of algorithms but operates within their framework" [132].

Algorithmic harm workarounds. Sometimes people act to mitigate individual algorithmic harms by avoiding them, creating distance between themselves and potential harm (e.g., [104, 124]). For example, Chinese food delivery workers chose not to follow platform-suggested routes, instead relying on their own knowledge of faster shortcuts [147]. In another instance, Bangladeshi users who felt misunderstood and unreasonably censored by Western Facebook moderation policies changed the ways that they posted—adding English translations, combining Bengali and English letters to write, and refraining from posting publicly [115].

Avoidance can also occur before harmful algorithmic behavior presents. A Google Images user anticipated a heterosexual bias in "wedding" search results and preemptively switched to use the search term "lesbian wedding" from the start [45]. As another example, Amazon began using AI Netradyne cameras that frequently punished delivery drivers for unsafe driving in situations that were not unsafe or not the driver's fault. Some drivers responded by driving over cautiously to preemptively avoid punishment, such as braking "once before a stop sign for the Netradyne camera, and another time for visibility before crossing an intersection" [65].

Instead of changing their usage to operate around harms, some actors protect themselves by masking algorithmic harm from their view. This can present in the form of using built-in platform affordances to block certain content or people [45]. For example, TikTok users pressed "not interested" buttons associated with content they

did not wish to see or found violating [123], and Bangladeshi Facebook users who saw problematic advertisements responded by unfollowing related pages as well as muting ads using the platform's settings [117]. Other times, this masking is less supported by platforms and their affordances. Ad blocking is a prevalent example of this, wherein users who dislike the experiences they receive from advertising algorithms use third-party blockers to remove the issues from their experiences [94]. TikTok users also sometimes protect themselves by intentionally seeking out and engaging more with videos similar to what they want to see, in order to shape algorithmic behavior [123].

Beyond self-protection. The workarounds above tend to protect only the person responding. In the above examples, others can be impacted secondarily, but these effects can be difficult to identify due to their diffuse nature. For example, if enough TikTok or Facebook users mark a specific post as uninteresting, the recommendation algorithm might deprioritize that video preemptively on other users' pages.

People also directly act to mitigate individual algorithmic harms that can protect many others beyond those taking action. For example, TikTok users who believed that the For You Page algorithm suppressed videos with social-justice content responded by engaging with apparently suppressed videos en masse: users commented, liked, and shared these videos repeatedly in attempts to boost video content that might have otherwise been hidden [79]. These actions aimed to amplify content broadly on the platform, using the For You Page algorithm to spread potentially previously suppressed content to many other TikTok users. And social workers using an AI tool individually adjusted and compensated for biased and incorrect suggested child-maltreatment risk scores, with the goal of protecting children and families [35, 80]. Similarly, Indian community health workers re-ran tests when they believed an AI app's diagnosis was incorrect [97].

Other times, actions to mitigate algorithmic harms may protect a subset of users. For instance, in response to beliefs that the For You Page algorithm suppressed videos created by people of certain identities, such as LGBTQ+ and Black creators, some TikTok content creators with more algorithmically favored identities shared their accounts with suppressed creators, allowing amplification of algorithmically suppressed voices via new platforms [79]. Additionally, some TikTok creators whose videos were suppressed responded by simply re-uploading videos that were taken down or by changing the aesthetics of their videos to better fit what they believed the algorithm prioritized [79, 123]. These actions aimed to amplify content mainly to the set of TikTok users who engaged with a certain content creator, either an algorithmically favored one in the former example or an algorithmically suppressed one in the latter ones.

4.3 Pursuing Legal Avenues Against Algorithmic Harm

In many cases, actors from below take legal action in response to encountering algorithmic harm. In this type of response, people attempt to leverage legal institutions to shift power away from algorithmlords; thus, we consider these responses to be a form of *pressuring*.

One way people do this is by suggesting and signing bill initiatives that aim to curb harmful algorithmic behaviors that they have encountered. For instance, three Californians, upset with how technology companies collect and use their personal data in algorithmic systems, proposed a ballot initiative for consumer privacy [139]. More than 600,000 other Californians signed in support of the initiative, certifying it so that it could later pass into law [138, 139].

People also exert pressure via direct legal action against algorithmlords. As one example, many business owners, believing that Yelp algorithmically removed positive reviews from their profiles because they did not pay for Yelp's advertising, filed almost 700 Federal Trade Commission reports against Yelp [52, 120]. Additionally, Amazon has recently been sued for "selling suicide kits", or algorithmically recommending buying combinations of items expedient for poisoning oneself, by parents of users who have died by this method [12, 13, 67].

Finally, due to the concentration and predominance of algorithmlords in colonialist (often Global North) countries, entire jurisdictions (often at some point colonized and/or Global South) occupy positions of subordinate power in and around algorithmic systems. As such, governments of these nations can respond from below to colonialist, exclusionary algorithmic systems by creating new policies to force algorithmlords to behave in ways less harmful to them. Indigenous Data Sovereignty is one example of this, advocating for "the right of Indigenous peoples to control data from and about their communities and lands, articulating both individual and collective rights to data access and to privacy" [100]. Indigenous groups worldwide have adopted processes and principles supporting this, such as the Māori in Aotearoa/New Zealand, the Sámi in Sweden, and the Ktunaxa Nation in Canada [102]. In another example case, much African population data must be stored in servers outside the continent since non-African companies own much of the data infrastructure; in response, African governmental leaders are instituting policies to make these companies adhere to privacy regulations and pay taxes, shifting some of the governance to themselves [10, 49, 55, 100, 133].

4.4 Investigating Potentially Harmful Algorithmic Behaviors Further

Another type of response from below to harmful algorithmic behaviors involves gathering more information about an issue. For instance, when Rotterdam's welfare fraud algorithm began unreasonably ranking people as high risk, one affected person requested the information contributing to their score [31]. This information can be leveraged to add legitimacy to people's claims and convince other groups with more power to listen and take action as well. As these investigations primarily help disempowered people build power, we view this response as *strengthening*.

Investigation can build people's power by helping them figure out what to do next and legitimize their claims. For example, San Diego residents dedicated significant effort to reading documents related to proposed smart streetlights as a "fact-finding mission" [142]. These people then leveraged what they had learned to refine their arguments and ensure use of language legible to those they engaged with [142].

Shen, DeVos et al. introduced the concept of "everyday algorithm auditing" to describe the process by which people "detect, understand, and/or interrogate problematic machine behaviors via their day-to-day interactions with algorithmic systems" [120]. For example, users conducting an everyday algorithm audit of potential racial bias in Twitter's image cropping algorithm carried out tests and gathered evidence around various hypotheses that they developed [120]. This helped users assess the extent to which racial biases were present and build up evidence that they used to push Twitter for change. As reviewed in [120], numerous cases of everyday algorithm audits have been documented in recent years.

4.5 Refusing Legitimate Engagement With Harmful Algorithmic Systems

In another type of response to algorithmic harm, people refuse to legitimately engage with harmful algorithmic systems. These responses draw on concepts of Indigenous and feminist refusal [15, 121] and especially parallel refusal of various technological practices [36, 61, 149, 150]. Thus, we consider these responses to be forms of *algorithmic refusal*.

Often people do this by removing some or all of themselves from an algorithmic system's visibility. This type of response can be *protective* in the sense that it requires people to employ their own power and, typically, limiting or eliminating true engagement serves to expose people to fewer harms, which makes their interactions safer. Additionally, this type of response can *pressure* by diminishing the engagement and data that algorithmic systems can consume. This penalizes those in charge of algorithmic systems and shifts some power and control to people acting from below.

People can also refuse legitimate engagement by actively seeking to impair or confuse algorithmic systems. These types of responses serve to *pressure* algorithmlords, making their disapproval of a system hard to ignore, and to *protect* themselves, obstructing an algorithmic system from its typical functioning.

4.5.1 Impairing an algorithmic system. Disempowered people sometimes react to harmful algorithmic systems by attempting to render them useless. For example, upset at being part of Waymo's self-driving car beta tests without their consent and at injuries these cars had caused, Arizonan pedestrians chose to fight back in response: they slashed the car tires and threw rocks at them [105, 107]. This protected the actors and others in the area to some extent, as it made self-driving cars less operable and sometimes forced humans inside the car to override the automated driving [107]. Additionally, this response pressured companies by making explicit the disapproval of potential customers. Similarly, when Amazon began using AI-powered cameras to surveil workers, some drivers covered the cameras with stickers so that they could no longer record, making the affected cameras inoperable [65].

In the same vein, people have put other potentially harmful algorithmic systems out of service. One way people do this involves disconnecting systems or making them otherwise unable to collect data. For example, a Carnegie Mellon University graduate student unplugged privacy-violating sensors that had been installed in his office without consent [64]. Similarly, before their exams, Chinese high-school students unplugged cameras used in facial recognition systems that monitored their behaviors during class [148]. The students had not consented and worried about accuracy, reliability, and bias within the system [148].

4.5.2 Depriving an algorithmic system. People sometimes refuse legitimate engagement with algorithmic systems by withholding needed inputs like their personal data or by removing those that have already been submitted. "Data strikes" in which "users withhold their data labor from a tech company, some of the company's essential services will suffer, and this would then force the company to make concessions that are desired by the public" constitute one significantly theorized about form of this response [136]. When done collectively, these responses aim to exert greater pressure but retain protective effects, whereas individually these responses protect more than they pressure.

Boycotts have been used for decades to pressure companies by refusing to engage, typically aiming to economically stunt via withholding money from companies. Boycotts operate similarly in digital contexts, often by removing user presence and thus revenue-generating ad impressions. For example, after Facebook disproportionately targeted Black users with misinformation ads, Facebook users boycotted the platform by logging out for a week [89, 140]. This put pressure on Facebook, but it also protected users by making them unable to encounter harmful Facebook ads for a time.

People can also leave harmful systems with similar aims. In one case, African workers at Sama, which performs some of Facebook's outsourced content moderation, left the company because of poor pay, poor working conditions, and poor mental health resulting from their job functions [98]. Resigning like this mainly serves to protect people from harmful work conditions and impacts, but it also mildly pressured Sama by showing that some were unwilling to work for them under current conditions.

People can also "leave" within the context of online algorithmic systems. For instance, in her "Opt Out" project Janet Vertesi says that deleting her Google account and ceasing use of Google-related tools "aimed at keeping [her] and [her] family away from the evils of data-collecting algorithmic systems" [134]. As another example, Twitter users frustrated by algorithmic control (and new ownership) responded by leaving Twitter and, in what has been called "conscious data contribution" [135, 137], moving to Mastodon, a platform without an algorithmic feed [127]. While these examples have underlying elements of pressure, they primarily serve to create safer existences online for those responding.

People can also withhold without the high commitment of entirely quitting all use of a platform, often via some form of hiding from an algorithmic system. For example, people adopt "practices of keeping personal and private aspects of life offline, such as not sharing images of young children and babies online, using encrypted services like Signal, [...and] the adoption of anonymous social media accounts to speak to a smaller circle of confidants" [60]. People can perform more manageable data strikes with smaller actions like using private browsers or refusing to rate, review, or comment on products [136]. For instance, Bangladeshi and Indian users, bothered by targeted ads, turned off location data or changing microphone settings for some apps on their phones to allow less access to their data [117]. Less severe actions like this lead to less pressure felt by companies or other targets. Responders gain some algorithmic anonymity and freedom from data collection, but the protection is

less complete than leaving a platform wholesale, with the trade-off of retaining the ability to interact with the system.

4.5.3 Confusing an algorithmic system. People also refuse legitimate engagement by providing false or unexpected information to harmful algorithmic systems. One version of this is "data poisoning", in which people provide "data that was created with the intention of thwarting the technology" [137]. Similarly, "obfuscation" encompasses "the deliberate addition of ambiguous, confusing, or misleading information to interfere with surveillance and data collection" [28].

This response often serves to obscure the personal data and behaviors of actors. Thus, confusing an algorithmic system can be protective. For example, Uber and Lyft drivers found their algorithmically determined ride payments inadequate, so they turned off their apps at coordinated times to feign a driver shortage and boost fares through app price surges [85, 130]. Baidu Deliveries, Ele.me, and Meituan delivery workers in China used similar methods, turning on work mode only at times or locations when the platform algorithms would not give them undesirable orders [129]. Acting individually, a user described pushing back against Facebook's advertising "data mining bot" by randomly and inaccurately liking pages [33]. Additionally, downloading and using some anti-tracking browser extensions enables this: for example, AdNauseam clicks ads automatically in the background to confuse online trackers [72]. These actions protect people from having their data collected and used harmfully while also subtly pressuring algorithmlords by forcing them to collect inaccurate data.

Confusing algorithmic systems can protect others beyond those responding. Obfuscating data has been described as a way to reduce algorithmic discrimination [57]. And Chinese online protesters created hundreds of duplicate videos of events by mirroring, cropping, and filtering them to evade algorithmic censorship [145]. As another example, after Twitter algorithmically amplified the racist hashtag "#whitelivesmatter", Twitter-using K-Pop fans "took over the hashtag [...] drowning out white-supremacist messages with nonsensical or anti-racist posts" [9]. Besides protecting, these responses pressure algorithmlords by rendering their algorithmic systems inoperable as intended.

4.6 Communicating Algorithmic Harm With Others

People also respond by communicating algorithmic harm with others to *protect, pressure,* or *strengthen.* This communication resembles past descriptions of users raising awareness about harms [45, 120]. These responses can involve communications from disempowered people to algorithmlords, to others affected by the same algorithmic system, or broadly to the public.

4.6.1 "Complaining" to those with more power. When people communicate to those with more power, like algorithmlords or members of legal institutions, they typically raise issues in hopes that those with more power will take action to address the conveyed harm. In doing this, people employ their existing power to try to create safer experiences for themselves while also putting a limited amount of pressure on algorithmlords. "Complaining" can have negative, patronizing connotations; we reference Ahmed's conception of complaints in which they can be powerful ways to change harmful structures, and we format this type of response in quotation marks to signal this [16].

Disempowered people have described submitting, flagging, or reporting issues as harmful when these options are built into platforms [45]. For example, YouTube creators harmed by demonetization errors in the advertising algorithm responded by contesting the errors using the platform's appeals system [83]. As another example, Wikipedia users surfaced algorithmic harms to Wikipedia platform developers, collaborating so that developers could then work to address the issues [66]. In these cases, people raise issues, using means often built into an algorithmic platform or otherwise systemically supported, to employ their own powers.

Sometimes existing pathways are nonexistent or insufficient, so people communicate in other ways. For example, after Facebook unreasonably restricted some Bangladeshi users' accounts, affected users failed to reach adequate resolution using the platform's appeals process; in response, several users asked people they knew who worked at Facebook for help restoring their accounts [115]. And workers at Sama, mentioned previously, sent management a list of their stipulations to improve their working conditions [98]. Voicing concerns to algorithmlords exerts light pressure: collaborative convincing to make changes, or beginning what could become a more involved and hostile process of shifting power through more intense, forceful means.

To exert more pressure on algorithmlords, people voice concerns about harmful algorithmic behaviors to legislators. For example, San Diego residents, troubled by potential surveillance from a smart streetlights program, showed up at community forums and lobbied city council members [142]. And a worker at Plastic Forte in Alicante, Spain, submitted a complaint to the Spanish data protection agency (AEDP) after discovering that the company monitored employee working times using secret facial recognition [37, 42].

4.6.2 Publicizing algorithmic harms broadly. When communicating issues broadly, people often aim to create bad PR for algorithmlords and connect with others who might help support their cause. Thus, publicizing can be viewed as both pressuring and strengthening. These two aims feed on each other, with increased media broadcasting leading to increased public awareness and support, which in turn often leads to more media coverage.

This type of response can be as straightforward as posting about encountered algorithmic harms on social media. For example, LGBTQ+ YouTubers harmed by automated video restrictions took to Twitter to describe the issues [6, 54, 120], as did lawyer Carrie Goldberg to share widely about Amazon's harmful "suicide kits" [12]. And Weibo users posted with the hashtag "#ThankGodIGraduatedAlready" to spread awareness about surveillance cameras installed in Chinese highschools [148]. This can lead to additional support, such as through retweets on Twitter, which builds users' power. If the posts gain significant popularity, the media often picks up the issues, which pressures algorithmlords.

People also publicize algorithmic harms broadly through protests. For example, when algorithmically determined A-level exam scores affected university admissions for students, especially those from lower socio-economic backgrounds, in over 160 countries including the UK and Bangladesh, students and teachers protested and

demonstrated worldwide [8, 48, 101, 118]. This garnered significant public support and many news outlets covered the protests, both strengthening those affected and pressuring algorithmlords.

People also directly contact news outlets to publicize algorithmic harms they encountered. For example, San Diego smart streetlights opposers planned a press conference to disseminate their concerns [142]. And after Johanna Burai saw only light skin in image searches for hands, she created a website and ran a media campaign to spread awareness and create more links to her site, pushing it higher in Google results [30, 132]. Directly contacting media can shift power by exerting PR pressure, and the resultant publications can help the public gain legitimacy and thus build power.

Another way people publicize algorithmic harm broadly involves writing longer, often more formal pieces or reports. This can further legitimize their own arguments to the broader public, strengthening their position, and provide resources external to the system for others. Additionally, these writings exert pressure when they are shared and spread widely. For example, the smart streetlights opposers created a policy report, which they also publicized by sharing its findings with journalists [142]. And after researcher Joy Buolamwini's noticed facial recognition could only detect her face with a white mask, she wrote and published a research paper about the issues [4, 29].

4.6.3 Sharing, discussing, & organizing with others affected by the algorithmic system. People also respond by sharing information with others affected by the same algorithmic system. This communication can aid in the formation of counterpublic spaces, or "parallel discursive arenas" [56], like those in which users collaborate and discuss potentially harmful algorithmic behaviors [120]. These spaces help similarly affected people strengthen their positions, exchanging useful knowledge and building power in numbers, and protect themselves via the development of solidarity, which can actively make algorithmic spaces safer for them.

Sometimes people simply convey information to others. For example, when Booking.com users noticed the rating algorithm calculated higher overall ratings for hotels than they intended, they warned others about this in the text portion of their reviews [51, 120]. As another example, TikTok users saw harmful content on their For You Pages and created videos of their own describing the problematic elements of that content to others [123]. In these ways, users alerted other users to harms present and tried to educate those who might not recognize the same harms.

Beyond simply sharing information, people also discuss issues with others affected. For example, Chinese riders for food delivery platforms created and joined small WeChat groups to develop networks for solidarity, sharing knowledge like difficult delivery locations [147]. And African workers at "ethical AI" company Sama created and used a WhatsApp group to discuss traumas of moderation work, low pay, lack of benefits, and unreasonable work hours [26, 98]. In these cases, people developed mutual support to protect themselves while discussing the nuances of an issue and potential directions for remediation, which helps them build power.

Finally, people also organize into collectives to take action with others, which can be used to protect themselves and pressure algorithmlords. For instance, smart streetlight adversaries in San Diego formed a coalition called TRUST San Diego to work together to counter the smart streetlights program [142]. The Chinese food delivery workers mentioned above used their chat groups for mutual aid amongst themselves, supporting each other in ways like informal transfers of orders to get deliveries completed on time and the formation of unofficial unions that provide increased bargaining power [26, 129, 147]. And African content moderation workers for ChatGPT, TikTok, and Facebook formed the African Content Moderators Union, strengthening themselves to push for better pay and work conditions [99].

5 LIMITATIONS

Though we attempted to ensure that the responses in our taxonomy are drawn from cases around the world, an early overrepresentation of Global North cases suggests a bias in the field that impacts our work despite our efforts. Regardless, differences between observed cultures and their norms of behavior, between designers' values and the relevant actors' values, and between pertinent legal systems can drastically affect how extensible a response is from one context to another. For example, a response of public protest that works well in one place might be much more dangerous or less efficacious in another. Additionally, we did not try to collect undocumented responses, so our analysis and taxonomy tends toward types of responses that receive some amount of visibility or leave some trace. Future work should explore responses from below in additional contexts and cultures, as well as in other languages and formats. Further, while our broad scope allowed development of an overarching understanding of responses from below, it necessarily abstracted away many specifics of individual cases, such as different actors' affiliations, incentives, possible actions, and complex and diverse relationships with power like their positions in the matrix of domination [69].

6 DISCUSSION

We have presented ways that people currently respond from below to algorithmic harm, through a taxonomy that connects different types of responses to existing theorizations of power. In this section, we discuss significant challenges related to supporting responses from below and highlight open opportunity spaces for future research to better support disempowered actors and their responses to algorithmic harm.

6.1 Burdens of Responding to Algorithmic Harms From Below

While our work has shown the potential and power in driving toward positive change from below, it is critical to consider the burdens this can have on the actors, particularly members of marginalized communities. In our work, we observed that documented responses to algorithmic harms largely come from communities affected by algorithmic harm (e.g., people of color, sexual and gender minorities). Thus more burdens are thrust onto members of marginalized and underserved groups who are already more prone to be exposed to and negatively affected by societal and algorithmic harms [43].

These burdens can further disempower people who respond from below. When helping other actors like industry practitioners

to address algorithmic harm, their efforts tend to recede from view. In addition, many types of responses burden responders with the work of transforming and transporting knowledge between cultures, known as translation work, of some kind [96]. For instance, translation work can aim to convey the legitimacy of information about algorithmic harms to those with greater power [128]. And efforts to support the disempowered in responding to algorithmic harm can simultaneously serve to prop up existing systems of power. For example, Turkopticon was intended to support Amazon Mechanical Turk workers, but its public legacy emphasized the value of design rather than the value of labor or laborers [73].

Ideally everyday people's concerns would be valued by default, but given the reality we see translation work as invaluable, especially when governmental institutions and actors from below are not in sync and communication in comprehensible formats is crucial for alignment. The FAccT community can collaborate with everyday people affected by algorithmic systems to build tools designed to push algorithmic complaints into view in ways that cannot be ignored. Building on existing repositories that collect and document real-world algorithmic harms (e.g., [1, 2]), platforms like this could be developed with and for everyday people, focusing on them as the primary users and contributors, and pressure algorithmlords by bringing issues to the attention of policymakers and the broader public.

To lessen the burdens of translation work on everyday people, we see opportunities for the FAccT community both to help actors from below most effectively frame findings from specific investigations and to proactively shape what forms of knowledge and evidence are understood as legitimate by powerful actors, such as algorithmlords and policymakers. The FAccT community has opportunities to elevate people's complaints to ensure that they are heard and acted on, recasting what affected people have already been saying into forms that will be appropriately valued by those with more power [96, 128]. In this vein, Ahmed breaks down the labor behind making complaints legible to an institution [16].

That being said, affected people should not have the burden of duty to surface and address issues as they arise; ideally, algorithmlords should preemptively address algorithmic harms so that affected people are not forced to take action. Changes after harm has occurred can be insufficient: replacing algorithmically generated A-level scores after protests, as described in Section 4.6.1, still left an "algorithmic imprint" of different scores than had the algorithm never existed [48]. With policymakers, the FAccT community could support actors from below by developing new systems and structures that incentivize algorithmlords to proactively identify, recognize, and address issues, and especially holding them accountable to harms they put into production [137]. For example, the US Federal Trade Commission recently penalized a company with a harmful algorithmic system, forcing them to pay a fine and delete data and related AI models [27, 81].

6.2 (Adverse) Impacts of Responses From Below to Algorithmic Harm

Responses from below can significantly help remediate algorithmic harms, but they can also have less desirable effects. For instance, broadly publicizing Amazon's harmful "suicide kit" recommendations [12] also makes more people aware of this suicide method. Such consequences become even more challenging with the presence of bad actors who misuse responses, intentionally or not, to algorithmic harms.

It is difficult to fully anticipate what sorts of impacts a response from below might have down the line, especially given how expansive they can be. In our research we saw these responses have both individual and more widespread impacts. Additionally, although we focused on more immediate effects of responses in this paper, responses can also have significant indirect effects.

While there might not be well-defined ways to identify and eliminate the adverse impacts, designers and developers who aim to empower responses from below to algorithmic harms can make harmful misuse more difficult. For example, Fawkes is a researcher-developed system that allows users to protect their digital images from unwanted facial recognition by making visually imperceptible pixel-level changes [116]. In theory, someone could use Fawkes to impersonate someone else. But because Fawkes was designed for avoiding identification in *unauthorized* models, for which most people want to avoid being recognized, being mistakenly identified contributes to the collective goal of obfuscation without causing harm to individuals.. Additionally, to further retain focus on mitigation of algorithmic harms, designers could create tools aligned with principles of restorative and transformative justice, making them more difficult to redirect toward harm [110, 144].

The FAccT community might also further consider concepts of consent and agency. We see opportunities for researchers to investigate, understand, and work toward a world in which consent and refusal are equally viable options [22, 122]. Zong and Matias highlight the ways in which refusal can preemptively avert harm [150]; to that end, the FAccT community should also consider their own agency and when it might be most beneficial to not design, to not deploy, or to declare insufficient and retract [21, 44, 70].

6.3 Obstacles Faced When Responding From Below to Algorithmic Harms

The specific contexts of algorithmic systems present unique challenges for responses from below to succeed. It can be challenging to pin down issues and productive directions forward due to the opaque nature of algorithmic systems and the emergent nature of algorithmic behavior in real-world use contexts [32, 58]. Additionally, it can be unclear whom to hold accountable and whom to contact for recourse, due to the diffuse and debated nature of responsibility for algorithmic behaviors [88]. Furthermore, different groups of people have different values around algorithmic systems [75], which can impact whether and how they respond. For some, the dynamics of power around these systems create a view of AI as necessarily authoritative and correct, which can lead them to take on responsibility for system harms themselves and to abstain from responding to address these harms [78, 104]. And finally, for some, entire categories of responses may be infeasible. For example, one response we observed involves leaving an algorithmic system—which is infeasible for some, such as rideshare drivers whose livelihoods are attached to the system. As highlighted by such cases, people often lack the power to fully disengage from algorithmic systems [68].

Algorithmic contract. Similar to theories of a social contract [91, 108], the public has found themselves within what we call an *"algorithmic contract"* in which they have theoretically consented and chosen to trade information about themselves and their behaviors in exchange for the computational promise and benefits of algorithmic systems. Although this algorithmic contract purports to be an agreement between everyone, in reality it is difficult for the public to disengage from algorithmic systems; more realistically, the algorithmic contract is an agreement between technological giants, in which algorithmlords with power to determine which algorithmic systems to use, how to use them, and where to deploy them have consented to these algorithmic systems for the rest of society. This parallels ideas of a technocracy, in which experts are the ruling power [53, 63], and leads to the current situation, one rife with algorithmic harms [109]. When people attempt to take back some control over the algorithmic systems that affect their lives via responding from below to algorithmic harms, they do so with few legal rights.

The FAccT community can work to improve legal protections and rights, and thus power, of those responding from below to algorithmic harm. We see researchers as having a responsibility to intervene and participate in the development of policy, helping create needed regulation that can have concrete positive impact in specific, real-world situations [126]. For instance, researchers could push to extend rulings like those recently won by Sandvig and the ACLU, which increased protections for researchers and journalists who break website terms of service to audit potentially harmful systems [11], to cover the public and their investigations. Researchers can also work with policymakers to create regulations that help people leave harmful systems. For instance, the EU's General Data Protection Regulation (GDPR) requires that users consent to data collection and stipulates users' rights to have their personal information erased [5]. Laws like this start to make it possible for people to fully separate from algorithmic systems and are especially important to actors for whom leaving an algorithmic system might be riskier. Additionally, researchers could develop and lead more workshops, special interest groups, or other discussions (e.g., [14, 86, 146]) that focus on ways that policy can support people's responses to algorithmic harm or render the responses unnecessary in the first place. FAccT community members can also serve as expert advisors providing oversight and nuance for legal pathways that shift power to help ensure that they sufficiently remediate algorithmic harm.

We see opportunities for the FAccT community to reframe the idea that algorithmlords or others with systemic power and expertise retain control throughout the deployment of an algorithmic system. The development of new forms of governance of algorithmic systems could allow for everyday people to be involved and empowered from the start, *by design*.

7 CONCLUSION

In this paper, we characterize real-world cases of responses from below to algorithmic harm and synthesize these cases into a taxonomy that categorizes responses based on their behaviors and their power-related aims. In doing so, our taxonomy offers the FAccT community (1) a set of major types of responses from below that future work can aim to support; and (2) a synthesis of existing theories on power and actions from below in and around algorithmic systems from fields such as anthropology, HCI, and communication. We call on FAccT community members to actively work with affected people to empower their efforts from below to safeguard themselves from algorithmic harm, and we present several existing opportunities for this.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) under Award No.2040942, The Jacobs Foundation, and a Cisco Research Award. We thank Franky Spektor, Jonathan Zong, Charvi Rastogi, Nathan DeVrio, and our anonymous reviewers for helping us improve our work.

REFERENCES

[1] [n. d.]. About. https://incidentdatabase.ai/about/
[2] [n. d.]. About the AIAAIC Repository. https://www.aiaaic.org/aiaaic-repository/about-the-aiaaic-repository
[3] [n. d.]. ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT). https://facctconference.org/
[4] [n. d.]. The Algorithmic Justice League: Mission, Team and Story. https://www.ajl.org/about
[5] 2016. Regulation (EU) 2016/679 (General Data Protection Regulation).
[6] 2017. YouTube changes restrictions on gay-themed content following outcry. https://www.theguardian.com/music/2017/mar/21/youtube-changes-restrictions-gay-lgbtq-themed-content-tegan-sarah
[7] 2019. We can't let Facebook do this again. https://naacpculpeper.org/we-cant-let-facebook-do-this-again/
[8] 2020. A-levels and gcses: How did the exam algorithm work? https://www.bbc.com/news/explainers-53807730
[9] 2020. K-pop fans take over #whitelivesmatter hashtag. https://www.nbcnews.com/news/asian-america/k-pop-fans-take-over-whitelivesmatter-hashtag-n1223376
[10] 2020. Who we are: The beginning of a new era. https://smartafrica.org/who-we-are/
[11] 2021. Statement on Supreme Court Decision Removing Hurdles to Online Civil Rights Testing and Research. https://www.aclu.org/press-releases/statement-supreme-court-decision-removing-hurdles-online-civil-rights-testing-and
[12] 2022. Another wrongful death lawsuit from C.A. Goldberg, PLLC against Amazon.com, Inc. https://www.cagoldberglaw.com/wrongful-death-plaintiffs-vs-amazon-com-inc-c-a-goldberg/
[13] 2022. Twitter thread. https://twitter.com/cagoldberglaw/status/1578121292502409216?s=20&t=prggAK5nTq-Xc2IwuRfICA Last Friday, CBS cancelled a segment about our clients suing Amazon for selling suicide kits to their now deceased kids. CBS' cowardice gave me renewed clarity about how urgent this litigation is. 1/.
[14] 2023. SIGCHI's Involvement in Public Policy. https://sigchi.org/public-policy/
[15] Sara Ahmed. 2017. No. https://feministkilljoys.com/2017/06/30/no/
[16] Sara Ahmed. 2021. *Complaint!* Duke University Press. http://www.jstor.org/stable/j.ctv1v7zdh2
[17] Ali Alkhatib. 2021. To Live in Their Utopia: Why Algorithmic Systems Create Absurd Outcomes. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 95, 9 pages. https://doi.org/10.1145/3411764.3445740
[18] Amy Allen. 2022. Feminist Perspectives on Power. In *The Stanford Encyclopedia of Philosophy* (Fall 2022 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University.
[19] Micah Altman, Alexandra Wood, and Effy Vayena. 2018. A Harm-Reduction Framework for Algorithmic Fairness. *IEEE Security & Privacy* 16, 3 (2018), 34–45. https://doi.org/10.1109/MSP.2018.2701149
[20] Solon Barocas, Kate Crawford, Aaron Shapiro, and Hanna Wallach. 2017. The problem with bias: Allocative versus representational harms in machine learning. In *9th Annual conference of the special interest group for computing, information and society*.
[21] Eric P.S. Baumer and M. Six Silberman. 2011. When the Implication is Not to Design (Technology). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 2271–2274. https://doi.org/10.1145/1978942.1979275

[22] Ruha Benjamin. 2016. Informed Refusal: Toward a Justice-based Bioethics. *Science, Technology, & Human Values* 41, 6 (2016), 967–990. https://doi.org/10.1177/0162243916656059 arXiv:https://doi.org/10.1177/0162243916656059

[23] W. Lance Bennett and Alexandra Segerberg. 2012. THE LOGIC OF CONNECTIVE ACTION. *Information, Communication & Society* 15, 5 (2012), 739–768. https://doi.org/10.1080/1369118X.2012.670661 arXiv:https://doi.org/10.1080/1369118X.2012.670661

[24] Abeba Birhane. 2021. Algorithmic injustice: a relational ethics approach. *Patterns* 2, 2 (2021), 100205. https://doi.org/10.1016/j.patter.2021.100205

[25] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. 2020. Language (Technology) is Power: A Critical Survey of "Bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485

[26] Masha Borak. 2022. China's Gig Workers are Challenging Their Algorithmic Bosses. https://www.wired.com/story/chinas-gig-workers-challenging-algorithmic-bosses/

[27] Ben Brody. 2022. Weight Watchers must delete algorithms built from kids' data. https://www.protocol.com/bulletins/weight-watchers-coppa-ftc

[28] Finn Brunton and Helen Nissenbaum. 2015. *Obfuscation: A User's Guide for Privacy and Protest.* The MIT Press.

[29] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (Proceedings of Machine Learning Research, Vol. 81)*, Sorelle A. Friedler and Christo Wilson (Eds.). PMLR, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

[30] Johanna Burai. 2015. WORLD WHITE WEB. https://johannaburai.com/World-White-Web

[31] Matt Burgess, Evaline Schot, and Gabriel Geiger. 2023. This Algorithm Could Ruin Your Life. https://www.wired.com/story/welfare-algorithms-discrimination/

[32] Stephen Bush. 2022. Beware the rise of the black box algorithm. https://www.ft.com/content/3d5556c5-520e-497a-aa5e-2546c5bc50cf

[33] Max C. 2011. Unsell Yourself - A Protest Model Against Facebook. https://yalelawtech.org/2011/05/10/unsell-yourself-a-protest-model-against-facebook/

[34] Stevie Chancellor, Michael L. Birnbaum, Eric D. Caine, Vincent M. B. Silenzio, and Munmun De Choudhury. 2019. A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA) *(FAT* '19)*. Association for Computing Machinery, New York, NY, USA, 79–88. https://doi.org/10.1145/3287560.3287587

[35] Hao-Fei Cheng, Logan Stapleton, Anna Kawakami, Venkatesh Sivaraman, Yanghuidi Cheng, Diana Qing, Adam Perer, Kenneth Holstein, Zhiwei Steven Wu, and Haiyi Zhu. 2022. How Child Welfare Workers Reduce Racial Disparities in Algorithmic Decisions. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 162, 22 pages. https://doi.org/10.1145/3491102.3501831

[36] Marika Cifor, Patricia Garcia, T.L. Cowan, Jasmine Rault, Tonia Sutherland, Anita Say Chan, Jennifer Rode, Anna Lauren Hoffmann, Niloufar Salehi, Lisa Nakamura, and et al. 2019. Feminist Data Manifest-No. https://www.manifestno.com/

[37] Karen Clayton. 2023. A company tracked its employees using facial recognition. https://thenationview.com/world-news/180989.html

[38] Eric Corbett, Emily Denton, and Sheena Erete. 2023. Power and Public Participation in AI. In *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (Boston, MA, USA) *(EAAMO '23)*. Association for Computing Machinery, New York, NY, USA, Article 8, 13 pages. https://doi.org/10.1145/3617694.3623228

[39] Rachel Courtland. 2018. Bias detectives: The researchers striving to make algorithms fair. https://www.nature.com/articles/d41586-018-05469-3

[40] Kate Crawford. 2017. The trouble with bias. In *Conference on Neural Information Processing Systems, invited speaker.*

[41] Mack DeGeurin. 2022. Welp, there goes twitter's ethical AI team, among others as employees post final messages. https://gizmodo.com/twitter-layoffs-elon-musk-ai-ethics-1849743051

[42] Carlos del Castillo. 2023. Multada una fábrica de Alicante por hacer reconocimiento facial a sus empleados sin avisarles. https://www.eldiario.es/tecnologia/multada-fabrica-alicante-reconocimiento-facial-empleados-avisarles_1_10152700.html

[43] Wesley Hanwen Deng, Boyuan Guo, Alicia DeVrio, Hong Shen, Motahhare Eslami, and Kenneth Holstein. 2023. Understanding Practices, Challenges, and Opportunities for User-Engaged Algorithm Auditing in Industry Practice. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 377, 18 pages. https://doi.org/10.1145/3544548.3581026

[44] Melissa Densmore. 2012. Claim Mobile: When to Fail a Technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1833–1842. https://doi.org/10.1145/2207676.2208319

[45] Alicia DeVos, Aditi Dhabalia, Hong Shen, Kenneth Holstein, and Motahhare Eslami. 2022. Toward User-Driven Algorithm Auditing: Investigating Users' Strategies for Uncovering Harmful Algorithmic Behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 626, 19 pages. https://doi.org/10.1145/3491102.3517441

[46] Peter Drahos. 1995. Information feudalism in the information society. *The Information Society - TIS* 11 (07 1995), 209–222. https://doi.org/10.1080/01972243.1995.9960193

[47] Veena Dubal. 2023. ON ALGORITHMIC WAGE DISCRIMINATION. *Columbia Law Review* 123, 7 (2023), pp. 1929–1992. https://www.jstor.org/stable/27264954

[48] Upol Ehsan, Ranjit Singh, Jacob Metcalf, and Mark Riedl. 2022. The Algorithmic Imprint. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1305–1317. https://doi.org/10.1145/3531146.3533186

[49] Nima Elmi. 2020. Is Big Tech Setting Africa Back? https://foreignpolicy.com/2020/11/11/is-big-tech-setting-africa-back/

[50] Sheena Erete, Yolanda Rankin, and Jakita Thomas. 2023. A Method to the Madness: Applying an Intersectional Analysis of Structural Oppression and Power in HCI and Design. *ACM Trans. Comput.-Hum. Interact.* 30, 2, Article 24 (apr 2023), 45 pages. https://doi.org/10.1145/3507695

[51] Motahhare Eslami, Kristen Vaccaro, Karrie Karahalios, and Kevin Hamilton. 2017. "Be Careful; Things Can Be Worse than They Appear": Understanding Biased Algorithms and Users' Behavior Around Them in Rating Platforms. *Proceedings of the International AAAI Conference on Web and Social Media* 11, 1 (May 2017), 62–71. https://doi.org/10.1609/icwsm.v11i1.14898

[52] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, and Karrie Karahalios. 2019. User Attitudes towards Algorithmic Opacity and Transparency in Online Reviewing Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300724

[53] JOHN H. EVANS. 2006. Between Technocracy and Democratic Legitimation: A Proposed Compromise Position for Common Morality Public Bioethics. *Journal of Medicine and Philosophy* 31, 3 (2006), 213–234. https://doi.org/10.1080/03605310600732834 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/03605310600732834 PMID: 16760101.

[54] Megan Farokhmanesh. 2018. YouTube is still restricting and demonetizing LGBT videos - and adding anti-LGBT ads to some. https://www.theverge.com/2018/6/4/17424472/youtube-lgbt-demonetization-ads-algorithm

[55] African Tax Administration Forum. 2020. ATAF PUBLISHES AN APPROACH TO TAXING THE DIGITAL ECONOMY. https://www.ataftax.org/ataf-publishes-an-approach-to-taxing-the-digital-economy

[56] Nancy Fraser. 1990. Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy. *Social Text* 25/26 (1990), 56–80. http://www.jstor.org/stable/466240

[57] Sorelle Friedler. 2017. Obfuscating Data to Prevent Discrimination. https://www.obfuscationworkshop.org/2017/10/obfuscating-data-to-prevent-discrimination/

[58] Batya Friedman and Helen Nissenbaum. 1996. Bias in Computer Systems. *ACM Trans. Inf. Syst.* 14, 3 (jul 1996), 330–347. https://doi.org/10.1145/230538.230561

[59] 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* (Chicago, IL, USA) *(FAccT '23)*. Association for Computing Machinery, New York, NY, USA, 205–216. https://doi.org/10.1145/3593013.3593989

[60] Maya Indira Ganesh and Emanuel Moss. 2022. Resistance and refusal to algorithmic harms: Varieties of 'knowledge projects'. *Media International Australia* 183, 1 (2022), 90–106. https://doi.org/10.1177/1329878X221076288 arXiv:https://doi.org/10.1177/1329878X221076288

[61] Patricia Garcia, Tonia Sutherland, Niloufar Salehi, Marika Cifor, and Anubha Singh. 2022. No! Re-Imagining Data Practices Through the Lens of Critical Refusal. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 315 (nov 2022), 20 pages. https://doi.org/10.1145/3557997

[62] Martin Giles. 2018. It's time to rein in the data barons. https://www.technologyreview.com/2018/06/19/240453/its-time-to-rein-in-the-data-barons/

[63] Bruce Gilley. 2017. Technocracy and democracy as spheres of justice in public policy. *Policy Sciences* 50, 1 (March 2017), 9–22. https://doi.org/10.1007/s11077-016-9260-2

[64] Eileen Guo and Tate Ryan-Mosley. 2023. Computer scientists designing the future can't agree on what privacy means. https://www.technologyreview.com/2023/04/03/1070665/cmu-university-privacy-battle-smart-building-sensors-mites/

[65] Lauren Kaori Gurley. 2021. Amazon's AI Cameras Are Punishing Drivers for Mistakes They Didn't Make. https://www.vice.com/en/article/88npjv/amazons-ai-cameras-are-punishing-drivers-for-mistakes-they-didnt-make

[66] Aaron Halfaker and R. Stuart Geiger. 2020. ORES: Lowering Barriers with Participatory Machine Learning in Wikipedia. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW2, Article 148 (oct 2020), 37 pages. https://doi.org/10.1145/3415219

[67] Joe Hernandez. 2022. A parents' lawsuit accuses Amazon of selling suicide kits to teenagers. https://www.npr.org/2022/10/09/1127686507/amazon-suicide-teenagers-poison

[68] Kashmir Hill. 2019. I Cut the 'Big Five' Tech Giants From My Life. It Was Hell. https://gizmodo.com/i-cut-the-big-five-tech-giants-from-my-life-it-was-hel-1831304194

[69] Patricia. Hill Collins. 2000. *Black feminist thought : knowledge, consciousness, and the politics of empowerment* (2nd edition. ed.). Routledge, New York.

[70] Sarah Homewood. 2019. Inaction as a Design Decision: Reflections on Not Designing Self-Tracking Tools for Menopause. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290607.3310430

[71] Lara Houston, Steven J. Jackson, Daniela K. Rosner, Syed Ishtiaque Ahmed, Meg Young, and Laewoo Kang. 2016. Values in Repair. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1403–1414. https://doi.org/10.1145/2858036.2858470

[72] Daniel C. Howe and Helen Nissenbaum. 2017. Engineering privacy and protest: A case study of AdNauseam. *CEUR Workshop Proceedings* 1873 (2017), 57–64. Funding Information: This publication has been supported in part by grants from US NSF CNS/NetS 105833, US NSF SATC 1642553, and the Research Grants Council of Hong Kong, China (Project No. CityU 11669616); 3rd International Workshop on Privacy Engineering, IWPE 2017 ; Conference date: 25-05-2017.

[73] Lilly C. Irani and M. Six Silberman. 2016. Stories We Tell About Labor: Turkopticon and the Trouble with "Design". In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4573–4586. https://doi.org/10.1145/2858036.2858592

[74] Steven J. Jackson and Laewoo Kang. 2014. Breakdown, Obsolescence and Reuse: HCI and the Art of Repair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 449–458. https://doi.org/10.1145/2556288.2557332

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

[76] Jakob Linaa Jensen. 2020. Digital Feudalism. In *The Medieval Internet: Power, Politics and Participation in the Digital Age*. Emerald Publishing Limited, 95–109. https://doi.org/10.1108/978-1-83909-412-520201008

[77] Pratyusha Kalluri. 2020. Don't ask if artificial intelligence is good or fair, ask how it shifts power. https://www.nature.com/articles/d41586-020-02003-2

[78] Shivani Kapania, Oliver Siy, Gabe Clapper, Azhagu Meena SP, and Nithya Sambasivan. 2022. "Because AI is 100% right and safe": User Attitudes and Sources of AI Authority in India. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 158, 18 pages. https://doi.org/10.1145/3491102.3517533

[79] Nadia Karizat, Dan Delmonaco, Motahhare Eslami, and Nazanin Andalibi. 2021. Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 305 (oct 2021), 44 pages. https://doi.org/10.1145/3476046

[80] Anna Kawakami, Venkatesh Sivaraman, Hao-Fei Cheng, Logan Stapleton, Yanghuidi Cheng, Diana Qing, Adam Perer, Zhiwei Steven Wu, Haiyi Zhu, and Kenneth Holstein. 2022. Improving Human-AI Partnerships in Child Welfare: Understanding Worker Practices, Challenges, and Desires for Algorithmic Decision Support. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 52, 18 pages. https://doi.org/10.1145/3491102.3517439

[81] Kate Kaye. 2022. The FTC's new enforcement weapon spells death for algorithms. https://www.protocol.com/policy/ftc-algorithm-destroy-data-privacy

[82] Os Keyes, Jevan Hutson, and Meredith Durbin. 2019. A Mulching Proposal: Analysing and Improving an Algorithmic System for Turning the Elderly into High-Nutrient Slurry. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290607.3310433

[83] Sara Kingsley, Proteeti Sinha, Clara Wang, Motahhare Eslami, and Jason I. Hong. 2022. "Give Everybody [..] a Little Bit More Equity": Content Creator Perspectives and Responses to the Algorithmic Demonetization of Content Associated with Disadvantaged Groups. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 424 (nov 2022), 37 pages. https://doi.org/10.1145/3555149

[84] Colin Koopman. 2017. The power thinker: Original, painstaking, sometimes frustrating and often dazzling. Foucault's work on power matters now more than ever. https://aeon.co/essays/why-foucaults-work-on-power-is-more-important-than-ever

[85] Bogdan Kulynych, Rebekah Overdorf, Carmela Troncoso, and Seda Gürses. 2020. POTs: Protective Optimization Technologies. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 177–188. https://doi.org/10.1145/3351095.3372853

[86] Jonathan Lazar. 2015. Public Policy and HCI: Making an Impact in the Future. *Interactions* 22, 5 (aug 2015), 69–71. https://doi.org/10.1145/2807916

[87] Christopher A. Le Dantec. 2016. Design through Collective Action / Collective Action through Design. *Interactions* 24, 1 (dec 2016), 24–30. https://doi.org/10.1145/3018005

[88] Theodore M. Lechterman. 2022. The Concept of Accountability in AI Ethics and Governance. In *The Oxford Handbook of AI Governance*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197579329.013.10

[89] Hanlin Li, Nicholas Vincent, Janice Tsai, Jofish Kaye, and Brent Hecht. 2019. How Do People Change Their Technology Use in Protest? Understanding. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 87 (nov 2019), 22 pages. https://doi.org/10.1145/3359189

[90] Gabriel Lima, Nina Grgić-Hlača, and Meeyoung Cha. 2023. Blaming Humans and Machines: What Shapes People's Reactions to Algorithmic Harm. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 372, 26 pages. https://doi.org/10.1145/3544548.3580953

[91] John Locke. 1960. *Two Treatises of Government*. The New American Library.

[92] Kristine Lu. 2021. Designing Democratic Systems for Civic Collective Action. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) *(CSCW '21)*. Association for Computing Machinery, New York, NY, USA, 270–274. https://doi.org/10.1145/3462204.3481792

[93] Kristian Lum and Rumman Chowdhury. 2021. What is an "algorithm"? it depends whom you ask. https://www.technologyreview.com/2021/02/26/1020007/what-is-an-algorithm/

[94] Arunesh Mathur, Jessica Vitak, Arvind Narayanan, and Marshini Chetty. 2018. Characterizing the Use of Browser-Based Blocking Extensions to Prevent Online Tracking. In *Proceedings of the Fourteenth USENIX Conference on Usable Privacy and Security* (Baltimore, MD, USA) *(SOUPS '18)*. USENIX Association, USA, 103–116.

[95] Milagros Miceli, Julian Posada, and Tianling Yang. 2022. Studying Up Machine Learning Data: Why Talk About Bias When We Mean Power? *Proc. ACM Hum.-Comput. Interact.* 6, GROUP, Article 34 (Jan 2022), 14 pages. https://doi.org/10.1145/3492853

[96] Michael Muller. 2004. HCI as Translation Work: How Translation Studies can Inform HCI Research and Practice. In *Presentations at the CHI 2004 Workshop on Reflective HCI (CHI '04)*. Association for Computing Machinery, New York, NY, USA. https://dominoweb.draco.res.ibm.com/reports/rc23318.pdf

[97] Chinasa T. Okolo, Srujana Kamath, Nicola Dell, and Aditya Vashistha. 2021. "It cannot do all of my work": Community Health Worker Perceptions of AI-Enabled Mobile Health Applications in Rural India. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 701, 20 pages. https://doi.org/10.1145/3411764.3445420

[98] Billy Perrigo. 2022. Inside Facebook's African Sweatshop. https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/

[99] Billy Perrigo. 2023. 150 African Workers for ChatGPT, TikTok and Facebook Vote to Unionize at Landmark Nairobi Meeting. https://time.com/6275995/chatgpt-facebook-african-workers-union/

[100] Marie-Therese Png. 2022. At the Tensions of South and North: Critical Roles of Global South Stakeholders in AI Governance. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1434–1445. https://doi.org/10.1145/3531146.3533200

[101] Jon Porter. 2020. UK ditches exam results generated by biased algorithm after student protests. https://www.theverge.com/2020/8/17/21372045/uk-a-level-results-algorithm-biased-coronavirus-covid-19-pandemic-university-applications

[102] Stephanie Carroll Rainie, Tahu Kukutai, Maggie Walter, Oscar Luis Figueroa-Rodríguez, Jennifer Walker, and Per Axelsson. 2019. *Indigenous data sovereignty.*

African Minds and the International Development Research Centre (IDRC), Cape Town and Ottawa, 300–319. https://hdl.handle.net/10289/12918 21.

[103] Inioluwa Deborah Raji and Joy Buolamwini. 2019. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (Honolulu, HI, USA) *(AIES '19)*. Association for Computing Machinery, New York, NY, USA, 429–435. https://doi.org/10.1145/3306618.3314244

[104] Divya Ramesh, Vaishnav Kameswaran, Ding Wang, and Nithya Sambasivan. 2022. How Platform-User Power Relations Shape Algorithmic Accountability: A Case Study of Instant Loan Platforms and Financially Stressed Users in India. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1917–1928. https://doi.org/10.1145/3531146.3533237

[105] Ryan Randazzo. 2018. A slashed tire, a pointed gun, bullies on the road: Why do Waymo self-driving vans get so much hate? https://www.azcentral.com/story/money/business/tech/2018/12/11/waymo-self-driving-vehicles-face-harassment-road-rage-phoenix-area/2198220002/

[106] Noopur Raval, Rida Qadri, Richmond Y. Wong, Tamara Kneese, and Alex Hanna. 2022. Considerations for Building Solidarity among Academic and Tech Workers: Thinking through Access, Positionality and Limits to Collective Action. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 153, 3 pages. https://doi.org/10.1145/3491101.3516511

[107] Simon Romero. 2018. Wielding Rocks and Knives, Arizonans Attack Self-Driving Cars. https://www.nytimes.com/2018/12/31/us/waymo-self-driving-cars-arizona-attacks.html

[108] Jean-Jacques Rousseau. 1967. *The Social Contract and Discourse on the Origin of Inequality*. Washington Square Press.

[109] Henrik Skaug Sætra. 2020. A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government. *Technology in Society* 62 (2020), 101283. https://doi.org/10.1016/j.techsoc.2020.101283

[110] Niloufar Salehi, Roya Pakzad, Nazita Lajevardi, and Mariam Asad. 2023. Sustained Harm Over Time and Space Limits the External Function of Online Counterpublics for American Muslims. 6, CSCW1, Article 93 (apr 2023), 24 pages. https://doi.org/10.1145/3579526

[111] 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, Article 368 (oct 2021), 33 pages. https://doi.org/10.1145/3479512

[112] Hanna Schneider, Malin Eiband, Daniel Ullrich, and Andreas Butz. 2018. Empowerment in HCI - A Survey and Framework. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173818

[113] James C. Scott. 1985. *Weapons of the Weak: Everyday Forms of Peasant Resistance*. Yale University Press. xv–xxii pages. http://www.jstor.org/stable/j.ctt1nq836

[114] Bryan Semaan. 2019. 'Routine Infrastructuring' as 'Building Everyday Resilience with Technology': When Disruption Becomes Ordinary. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 73 (nov 2019), 24 pages. https://doi.org/10.1145/3359175

[115] Farhana Shahid and Aditya Vashistha. 2023. Decolonizing Content Moderation: Does Uniform Global Community Standard Resemble Utopian Equality or Western Power Hegemony?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 391, 18 pages. https://doi.org/10.1145/3544548.3581538

[116] Shawn Shan, Emily Wenger, Jiayun Zhang, Huiying Li, Haitao Zheng, and Ben Y. Zhao. 2020. Fawkes: Protecting Privacy against Unauthorized Deep Learning Models. In *29th USENIX Security Symposium (USENIX Security 20)*. USENIX Association, 1589–1604. https://www.usenix.org/conference/usenixsecurity20/presentation/shan

[117] Tanusree Sharma, Smirity Kaushik, Yaman Yu, Syed Ishtiaque Ahmed, and Yang Wang. 2023. User Perceptions and Experiences of Targeted Ads on Social Media Platforms: Learning from Bangladesh and India. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 663, 15 pages. https://doi.org/10.1145/3544548.3581498

[118] Sam Shead. 2020. How a computer algorithm caused a grading crisis in British schools. https://www.cnbc.com/2020/08/21/computer-algorithm-caused-a-grading-crisis-in-british-schools.html

[119] Renee Shelby, Shalaleh Rismani, Kathryn Henne, AJung Moon, Negar Rostamzadeh, Paul Nicholas, N'Mah Yilla, Jess Gallegos, Andrew Smart, Emilio Garcia, and Gurleen Virk. 2022. Sociotechnical Harms: Scoping a Taxonomy for Harm Reduction. https://doi.org/10.48550/ARXIV.2210.05791

[120] Hong Shen, Alicia DeVos, Motahhare Eslami, and Kenneth Holstein. 2021. Everyday Algorithm Auditing: Understanding the Power of Everyday Users in Surfacing Harmful Algorithmic Behaviors. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 433 (oct 2021), 29 pages. https://doi.org/10.1145/3479577

[121] Audra Simpson. 2007. On Ethnographic Refusal: Indigeneity, 'Voice' and Colonial Citizenship. *Junctures: The Journal for Thematic Dialogue* (2007), 67–80.

[122] Audra Simpson. 2017. The ruse of consent and the anatomy of 'refusal': cases from indigenous North America and Australia. *Postcolonial Studies* 20, 1 (2017), 18–33. https://doi.org/10.1080/13688790.2017.1334283 arXiv:https://doi.org/10.1080/13688790.2017.1334283

[123] Ellen Simpson and Bryan Semaan. 2021. For You, or For"You"? Everyday LGBTQ+ Encounters with TikTok. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3, Article 252 (jan 2021), 34 pages. https://doi.org/10.1145/3432951

[124] Anubha Singh and Tina Park. 2022. Automating Care: Online Food Delivery Work During the CoVID-19 Crisis in India. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 160–172. https://doi.org/10.1145/3531146.3533082

[125] Mona Sloane, Emanuel Moss, Olaitan Awomolo, and Laura Forlano. 2022. Participation Is not a Design Fix for Machine Learning. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (<conf-loc>, <city>Arlington</city>, <state>VA</state>, <country>USA</country>, </conf-loc>) *(EAAMO '22)*. Association for Computing Machinery, New York, NY, USA, Article 1, 6 pages. https://doi.org/10.1145/3551624.3555285

[126] Anne Spaa, Abigail Durrant, Chris Elsden, and John Vines. 2019. Understanding the Boundaries between Policymaking and HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300314

[127] Jaime Stathis. 2023. What Is Mastodon, and Why Are People Leaving Twitter for It? https://www.rd.com/article/what-is-mastodon/

[128] Cella M Sum, Anh-Ton Tran, Jessica Lin, Rachel Kuo, Cynthia L Bennett, Christina Harrington, and Sarah E Fox. 2023. Translation as (Re)Mediation: How Ethnic Community-Based Organizations Negotiate Legitimacy. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 603, 14 pages. https://doi.org/10.1145/3544548.3581280

[129] Ping Sun. 2019. Your order, their labor: An exploration of algorithms and laboring on food delivery platforms in China. *Chinese Journal of Communication* 12, 3 (2019), 308–323. https://doi.org/10.1080/17544750.2019.1583676 arXiv:https://doi.org/10.1080/17544750.2019.1583676

[130] Sam Sweeney. 2019. Uber, Lyft drivers manipulate fares at Reagan National causing artificial price surges. https://wjla.com/news/local/uber-and-lyft-drivers-fares-at-reagan-national

[131] Megan Twohey and Gabriel J.X. Dance. 2022. Lawmakers Press Amazon on Sales of Chemical Used in Suicides. https://www.nytimes.com/2022/02/04/technology/amazon-suicide-poison-preservative.html

[132] Julia Velkova and Anne Kaun. 2019. Algorithmic resistance: media practices and the politics of repair. *Information, Communication & Society* (2019), 1–18. https://doi.org/10.1080/1369118X.2019.1657162

[133] Quentin Velluet. 2021. Can Africa salvage its digital sovereignty? https://www.theafricareport.com/80606/can-africa-salvage-its-digital-sovereignty/

[134] Janet Vertesi. 2022. About the Opt Out Project. https://www.optoutproject.net/about-the-opt-out-project/

[135] Nicholas Vincent and Brent Hecht. 2021. Can "Conscious Data Contribution" Help Users to Exert "Data Leverage" Against Technology Companies? *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 103 (apr 2021), 23 pages. https://doi.org/10.1145/3449177

[136] Nicholas Vincent, Brent Hecht, and Shilad Sen. 2019. "Data Strikes": Evaluating the Effectiveness of a New Form of Collective Action Against Technology Companies. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 1931–1943. https://doi.org/10.1145/3308558.3313742

[137] Nicholas Vincent, Hanlin Li, Nicole Tilly, Stevie Chancellor, and Brent Hecht. 2021. Data Leverage: A Framework for Empowering the Public in Its Relationship with Technology Companies. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 215–227. https://doi.org/10.1145/3442188.3445885

[138] Daisuke Wakabayashi. 2018. California Passes Sweeping Law to Protect Online Privacy. https://www.nytimes.com/2018/06/28/technology/california-online-privacy-law.html

[139] Daisuke Wakabayashi. 2018. Silicon Valley Faces Regulatory Fight on Its Home Turf. https://www.nytimes.com/2018/05/13/business/california-data-privacy-ballot-measure.html

[140] Charlie Warzel. 2018. The NAACP Wants Users To Log Out Of Facebook In Protest. https://www.buzzfeednews.com/article/charliewarzel/naacp-log-out-

facebook-protest
[141] Laura Weidinger, Jonathan Uesato, Maribeth Rauh, Conor Griffin, Po-Sen Huang, John Mellor, Amelia Glaese, Myra Cheng, Borja Balle, Atoosa Kasirzadeh, Courtney Biles, Sasha Brown, Zac Kenton, Will Hawkins, Tom Stepleton, Abeba Birhane, Lisa Anne Hendricks, Laura Rimell, William Isaac, Julia Haas, Sean Legassick, Geoffrey Irving, and Iason Gabriel. 2022. Taxonomy of Risks Posed by Language Models. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 214–229. https://doi.org/10.1145/3531146.3533088
[142] Cedric Deslandes Whitney, Teresa Naval, Elizabeth Quepons, Simrandeep Singh, Steven R Rick, and Lilly Irani. 2021. HCI Tactics for Politics from Below: Meeting the Challenges of Smart Cities. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 297, 15 pages. https://doi.org/10.1145/3411764.3445314
[143] Yuxi Wu, W. Keith Edwards, and Sauvik Das. 2022. "A Reasonable Thing to Ask For": Towards a Unified Voice in Privacy Collective Action. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 32, 17 pages. https://doi.org/10.1145/3491102.3517467
[144] Sijia Xiao, Coye Cheshire, and Niloufar Salehi. 2022. Sensemaking, Support, Safety, Retribution, Transformation: A Restorative Justice Approach to Understanding Adolescents' Needs for Addressing Online Harm. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 146, 15 pages. https://doi.org/10.1145/3491102.3517614
[145] Damir Yalalov. 2022. Censorship AI algorithms: How Chinese users cheat them. https://mpost.io/censorship-ai-algorithms-how-chinese-users-cheat-them/
[146] Qian Yang, Richmond Y. Wong, Thomas Gilbert, Margaret D. Hagan, Steven Jackson, Sabine Junginger, and John Zimmerman. 2023. Designing Technology and Policy Simultaneously: Towards A Research Agenda and New Practice. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 343, 6 pages. https://doi.org/10.1145/3544549.3573827
[147] Zizheng Yu, Emiliano Treré, and Tiziano Bonini. 2022. The emergence of algorithmic solidarity: unveiling mutual aid practices and resistance among Chinese delivery workers. *Media International Australia* 183, 1 (2022), 107–123. https://doi.org/10.1177/1329878X221074793 arXiv:https://doi.org/10.1177/1329878X221074793
[148] Xue Yujie. 2019. Camera Above the Classroom. https://www.sixthtone.com/news/1003759
[149] Jonathan Zong. 2020. From Individual Consent to Collective Refusal: Changing Attitudes toward (Mis)Use of Personal Data. *XRDS* 27, 2 (dec 2020), 26–29. https://doi.org/10.1145/3433140
[150] Jonathan Zong and J. Nathan Matias. 2023. Data Refusal From Below: A Framework for Understanding, Evaluating, and Envisioning Refusal as Design. *ACM J. Responsib. Comput.* (oct 2023). https://doi.org/10.1145/3630107 Just Accepted.

A ADDITIONAL DETAILS ON PROCESS OF RESPONSE COLLECTION & CATEGORIZATION

To explore the space, we adopted a case study approach, gathering documented historical cases of responses from below that intend to remediate perceived algorithmic harm. We understand actors and actions from below as those lacking direct power to make decisions about the design and deployment of algorithmic systems. To emphasize power relations, we broadly refer to those who have control over algorithmic systems as *"algorithmlords"*, building on past work describing "data barons" who control data and "digital feudalism" in which tech companies behave as modern-day feudal lords [46, 62, 76]. In the vein of [142], we understand algorithmlords as those holding greater power to envision and construct algorithmic systems, occupying positions "above" those subjugated by their decisions, while acknowledging that relationships and interactions between people are more complex and nuanced than this fully captures. We included cases in which the system is described as algorithmic in the source material (making no attempt to further define what an algorithm is [93]), the actor perceives harm from that system and takes action in response, and the actor lacks direct power to make decisions about the design or deployment of algorithmic systems (following [142]).

We began with a set of 10 documented cases drawn from seven past research papers that the authors were familiar with [45, 51, 52, 79, 83, 120, 137]. From our initial set of research papers, we used both forward and backward citation chaining to find additional examples, following relevant citations to find other cases. We also conducted searches of research literature, news media, and social media using combinations of relevant keywords (e.g., "algorithmic", "harm", "users", "respond") to find additional relevant cases. Our cases spanned many different algorithmic domains (e.g., recommendation, image cropping, facial recognition) and types of harm (e.g., financial, stereotyping, surveillance).

While collecting cases, we also began analyzing and categorizing the responses present in the cases we had already gathered. We coded each case with the algorithmic harm identified by the responder(s), who was impacted, the algorithmic system, the actor(s) responding, the response action or actions, any tools used, and if described, the desired and actual outcomes. As a single case often included responding in multiple ways, we extracted all the responses from each case. We conducted a bottom-up thematic analysis with a series of interpretation sessions on these responses and, based on this, grouped our set of responses into *types* of responses based on similar responder behaviors present, which we then synthesized into three high-level *strategies*, each of which captures a different way that these responses aim to interact with power.

Finally, based on discussions with researchers beyond our team, we identified additional relevant cases that could inform the development and refinement of our taxonomy. For example, Global North locations were overrepresented in our initial set of cases. To mitigate this bias, we added specific responses from cases that emerged in our discussions and from conducting additional searches that covered other regions.

We conducted initial searches in the spring of 2022, then coded and interpreted as already described in the Appendix. We then conducted additional searches that targeted regions underrepresented in our dataset by appending region names to the previous search terms and identified an additional 50 responses across 26 cases, using the criteria described above. We iteratively integrated the cases into our categories, expanding one of our categories in the process. Continuing our searches in 2023, we identified 19 more responses over 11 cases that, when coded and integrated in our analysis, did not impact the taxonomy's structure.

In all, we collected 96 cases from 75 sources with 169 extracted responses from below to algorithmic harm. We stopped collecting cases when we reached saturation in our categorization: that is, when we stopped finding examples of new types of responses. We emphasize that we see the cases used as the basis for our taxonomy as illustrative of the phenomenon under study, but not necessarily exhaustive. Additionally, due to the constantly changing and emerging nature of algorithmic systems, of their applications, of related harms, and of related responses to these harms, we see this work as an initial exploration and invite future researchers to expand on it.

Received 22 January 2024; accepted 30 March 2024