NORTHWESTERN UNIVERSITY

Democratizing Power in Tech: Reconceptualizing Data Production as a Form of Labor

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Technology and Social Behavior

By

Hanlin Li

EVANSTON, ILLINOIS

December 2022

ABSTRACT

Democratizing Power in Tech: Reconceptualizing Data Production as a Form of Labor

Hanlin Li

Public-facing data-driven technologies such as social media platforms and search engines rely on data producers, such as users and crowd workers, to be feasible and financially sustainable. Recently, it became clear that the goals of these data-driven technologies do not always align with those of the public, causing public backlashes against such companies. Despite playing a crucial role in supporting data-driven technologies, data producers often do not have control over the downstream applications of the digital traces from their activities. My dissertation focuses on empowering data producers to leverage their data contribution to shape data practices and data-driven technologies. I do so by characterizing data production and identifying opportunities to recognize and support the unpaid labor data producers provide. In the first chapter, I examined how data producers protest prominent tech companies motivated by concerns about data monetization. In the second chapter, I construct a taxonomy of data labor to guide researchers, designers, and the general public in helping data producers gain control over the outcome of their data through collective means. The third chapter dives into a specific type of data

labor–content moderation on Reddit–and characterizes the invisibility of data producers' contribution to technology. The fourth chapter then focuses on quantifying the amount of data labor going in upholding Reddit communities. In the fifth chapter, I move onto the aspect of collective action in the data labor taxonomy and identify opportunities for collective means among data producers to gain leverage against technology companies. Finally, the sixth and final chapter extends the data labor taxonomy identifying broader directions and principles of data labor through a cross-disciplinary literature review of data governance, data markets, and worker-centered design work. Together, this dissertation characterizes data labor and charts out the path towards a data future in which data producers' interests and values shape the design and development of data-driven technologies.

ACKNOWLEDGEMENTS

I have been fortunate to receive tremendous support from a village of colleagues, mentors, friends, and family members over the years. They played a significant role in shaping my thinking and experience at Northwestern University and beyond. I will give a few shout-outs here but this list is by no means exhaustive.

First and foremost, I would like to thank my advisor Brent Hecht. Thank you for taking a chance with me, guiding me through our many many research discussions, and helping me become a better researcher and a better writer. As a non-native speaker with little training in academic writing, I struggled with completing a full paper myself in the early years of my graduate school. You not only offered thoughtful and encouraging feedback but also extended extensive resources. Throughout the years, I always looked forward to seeing your comments and annotations on my writing–which is perhaps an outlier among graduate students as illustrated in the popular comics [1], because I would always learn something new from you.

I would also like to thank my other advisor, Stevie Chancellor. I was very fortunate to have you join my village right before the pandemic. You made yourself available throughout our research collaborations and offered incredibly insightful feedback as I tackled various research challenges. Thank you for your generous support as I figured out my academic career path.

A huge thank you to my Master's advisor, Erin Brady. You inspired me to embark on this journey and showed me a research path that is driven by empathy and care. I cannot

[1]https://phdcomics.com/comics/archive.php?comicid=810

say thank you enough for your mentorship and unfailing support. Remember the book you gifted me as my graduation gift? Now I can finally pass it along to the next person as “Dr. Li”.

Thank you to the rest of my committee, all of whom have been my informal mentors during my time at Northwestern. Darren for answering all of my questions, from what research papers and methods I should consider to ”are these random mushrooms I found in my apartment courtyard morels?” Loren for pushing me to articulate my contribution and providing advice on my work and presentations over the years. Nick for encouraging me to improve the narrative of my work and refine my thinking.

Thanks to my research buddies, colleagues, and collaborators - Nick, Isaac, Allen, my peers from GroupLens, Jofish, Janice, Sarah, Bodhi, Lynn, Disha - for inspirations, guidance, and discussion. I learned so much from all of you.

Thanks to my parents and my grandparents, who supported my decision of living on the other side of the Pacific Ocean. Thanks to Ali and Kiwi for sticking by me and planning with me to venture out to the West Coast. Thanks to all my wonderful friends and support village - Jes, Alexa, Katya, Nicole, Yixue, Eureka, Yixiang, Ada, Ann, and many many others. Thanks to the Evanston Animal Shelter community and the dogs.

Table of Contents

Introduction

Prominent data-driven technologies such as social media platforms, predictive models, and search engines are made possible in part because data producers such as users and crowdworkers generate massive digital traces. For example, online communities like subreddits would not exist without community members and volunteer moderators actively sharing and managing content. Rating systems like Google Maps and Yelp would not be able to make recommendations without people producing ratings for businesses and locations. Put another way, many data-driven technologies are co-created by data producers and technology companies.

Despite data producers' crucial role in enabling and upholding these technologies, data-driven technologies are built without data producers' meaningful involvement and sometimes consent in the extraction and monetization of data. While in most cases data producers receive a free or low-cost service (e.g. Yelp or Facebook), it is unclear if the value proposition is symmetric. Technology companies freely privatize and monetize the data they collect and subsequently gain an immense amount of profit and power. Conversely, the vast majority of the data producer population have little to no decision-making power to shape how their digital traces are used and by whom, nor do they have the means to shape the powerful technologies they co-create with companies. This power imbalance becomes apparent when companies' goals do not align well with public interests and has led to public outcries around certain particularly harmful technologies. For example, gig work platforms have long been criticized for their lack of protection and support for their workers, whose service and data are fundamental for these platforms. More recently, companies that develop and deploy large language models have faced criticism for their misuse of user-generated content as training data and the potential risks associated with automatic content generation. More broadly, many platforms have been criticized for their prioritizing promoting user engagement over mitigating issues exacerbated by their platforms such as misinformation, hate speech, and social biases.

To mitigate the power imbalance between data-driven technologies and data producers, I seek to understand and characterize the latter's concerns and contributions to inform specific action plans and policy recommendations, before drawing from literature to envision principles for an alternative data future. My goal is to holistically empower data producers by providing recommendations about potential pathways toward developing data producer-centered technologies such as social platforms and advanced models. I first conducted a systematic survey of protests against technology companies to identify data producers' most prominent concerns about the tech industry. My study shows that concerns about data monetization has become one of the most prominent motivations for data producers to participate in a protest against technology companies, highlighting the importance of shifting our attention to data and its downstream impact in our study of human activities in computing systems.

Motivated by the public's interests in influencing technology companies' practices, in the second chapter, I develop a taxonomy of data labor that informs a roadmap for researchers, practitioners, and the general public to help data producers influence operators of data-driven technologies. In this work, I propose to reconceputalize data production as a form of labor–data labor–and lay out six prominent dimensions that characterize data labor. Drawing from the literature, I then prescribe collective action opportunities for all stakeholders of data-driven technologies to elevate data producers' status in their relationship with operators of data-driven technologies.

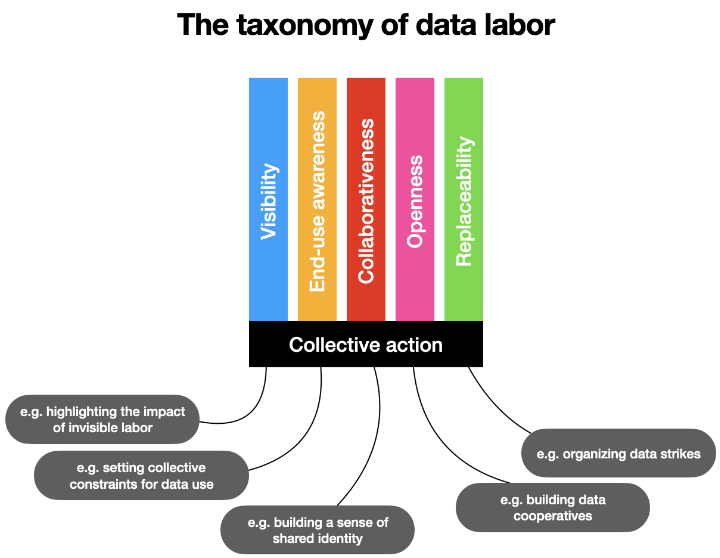


Figure 1. The taxonomy of data labor.
The taxonomy of data labor and how it informs opportunities for collective action.

In the third and fourth chapter, I dive into one specific dimension of data labor–visibility, because the lack of visibility about data labor's *material contribution* to technologies have been a key challenge in recognizing and empowering this labor. I focus on one specific instance of data labor, volunteer content moderation, and quantify the invisibility and value of this labor in Chapter Three and Four, respectively. These studies show that although data labor is often regarded as "free" and hidden by technology companies, it is possible to track and measure the monetary value of this labor subsidy. This work takes a step towards equipping data producers with knowledge about the value they bring into technology companies, and, thereby, assisting them in their negotiations with companies for support and resources such as better tools and better working conditions.

In the fifth chapter, I focus on the collective action aspect of the data labor taxonomy and explore design opportunities to assist data producers in realizing their power over data-driven technologies. I designed and deployed a browser extension that connects data producers alike and automate their protest actions (e.g. stop visiting a website). This study shows the potential for online collective action in protesting against technology companies while highlighting the challenges that future activists and researchers must contend with, such as lack of replacements for protest targets.

In the sixth and last chapter, I zoom out and draw from three bodies of related literature–data governance, data economics, and worker-centered design–to envision principles for ethical use of data, such as fair compensation, consent, and collective benefits. Notably, data producers are not considered as a stakeholder in the majority of data governance and data economics frameworks; when they are included, it is often in regard to mitigating privacy risks associated with data. Only recently, scholars started advocating for integrating data producer-centered values such as collective benefits. This chapter further outlines the future long-term research directions that will help to enforce these data producer-centered principles.

Overall, my dissertation shows how conceptualizing data production as labor is useful to inform immediate next steps and long-term plans on empowering the data-generating public to improve its status and influence over the technology industry.

Below, I use the pronoun "we" instead of "I" to reflect the collaborative nature of my work. I completed the vast majority of this dissertation with the support of my colleagues.

Chapter I: Protests against Data-Driven Technologies: the Status Quo

To understand issues prominent in data-driven technologies, I first sought to use surveys to understand public perceptions of prominent technology companies. Over the years, there have existed protests against such companies that involve people stopping or changing their use of these companies' products. These protests have attracted increasing public attention, including boycotts against Facebook to protest illicit data harvesting and the spread of misinformation [**Granville2018, Greenfield, Frier, and Brody2018**], boycotts against Uber to protest its behavior surrounding a taxi strike and sexual harassment in the company [**Semuels2017**], and boycotts against Amazon to protest working conditions and anti-tax lobbying [**Kasperkevic2018**].

Researchers have also become increasingly interested in these types of protests against technology companies [**Li et al.2018b, Matias2016b**]. For instance, Vincent et al. recently explored the concepts of "data strikes" and "data boycotts" against large-scale machine learning systems [**Vincent, Hecht, and Sen2019**], Posner and Weyl argued for the formation of "data unions" [**Posner and Weyl2019**] or other mediators of individual data [**Lanier and Weyl2018a**], and Li et al. developed technologies to scaffold these and other types of protests [**Li et al.2018b**]. More generally, Human-Computer Interaction (HCI) researchers [**Baumer2018a, Baumer et al.2015b, Satchell and Dourish2009, Wyatt, Oudshoorn, and Pinch2003**] have called for studying specific forms of non-use, of which recent protests against technology companies can be understood as a part.

However, despite the growing public and scholarly interest in protests against data-driven technologies, we lack critical empirical information about these protests. Core questions surrounding participation rates, tactics, and motivations remain unaddressed. Put another way, we do not know the extent of the population that is participating in one of these protests, nor do we have a rigorous understanding of their specific protest tactics or motivations for protesting. Additionally, we lack knowledge about what challenges people face in these protests and what roadblocks prevent people from protesting.

Through the results of two nationally-representative surveys, this paper contributes an improved descriptive understanding of whom we are calling protest users. These people are current or past users of a technology who change (protest use) or stop (protest non-use) their use of the technology due to the values or actions of the company behind the technology. Our surveys sampled adult Internet users in the United States. The first exploratory survey was conducted in 2017 (n = 463). The second survey was conducted in 2019 and directly targeted specific research questions about protest users (n = 398). In particular, we examined if, how, and why people have become protest users of five major technology (tech) companies (Amazon, Apple, Facebook, Google, Microsoft; five of the most valuable tech companies on the U.S. markets), and the challenges and roadblocks experienced by active and potential protest users, respectively.

Our results suggest that a surprisingly large share of web users in the United States are protest users. 30% of our 2019 respondents reported being active protest users of at least one tech company. This number is a meaningful increase from the 9% of respondents in our 2017 survey (although as we detail below, this comparison must be interpreted with caution). Furthermore, an additional 19% of our 2019 respondents who were not actively protesting expressed interest in doing so. In total (after rounding to the nearest percent), 48% of respondents indicated that they were either active or potential protest users.

Among active protest users, the most commonly reported motivations were concerns about business models that profit from user data and concerns about privacy (echoing previous findings about technology non-use and privacy concerns [**Baumer et al.2013, Stieger et al.2013, Young and Quan-Haase2013**]). Furthermore, stopping use entirely and using ad blockers were the most common tactics that our protest users reported employing against tech companies, and losing social connections was the most prevailing challenge protest users faced. Among our potential protest users, we observed that a major roadblock to protesting was a lack of alternative products. This finding is in alignment with current concerns around the monopoly power of technology companies and corresponding effects on the consumer's ability to shape company behavior [**Rogoff2019**]. We also observed some roadblocks that were especially prominent for particular companies. For instance, consistent with prior work [**Baumer et al.2013**], respondents reported that the possibility of "losing connections with others" and "missing out on information" prevented them from leaving Facebook.

From the lens of the literature on protests against technology companies, our study provides evidence that there could be substantial demand for technologies to support protest users and provides guidance for the design of these technologies. This guidance includes helping people protest collectively and aiding them in accessing alternative products and services. Our work also replicates some findings from the non-use literature (e.g. the importance of privacy concerns and demographic differences in non-use behavior) and identifies some characteristics of protest users that are unique relative to other types of non-use (e.g. motivations and tactics, specific demographic trends in protest non/use).

We begin below by covering work that inspired this research. We then discuss our survey methodology and results, before entering into our discussion of implications.

Related Work

In this section, we discuss the two literatures that most informed our overall thinking for this research: the technology non-use literature and the literature on protests against technology companies.

Technology Non-Use and "Non/Use"

Our ideation and study design for this project was influenced by the literature on non-use in science and technology studies (STS) and human-computer interaction (HCI) (e.g. [**Baumer et al.2013, Baumer et al.2015b, Hargittai2007, Satchell and Dourish2009, Schoenebeck2014, Wyatt, Oudshoorn, and Pinch2003**]). This body of work argues that, in contrast to prevailing perspectives in HCI, non-use can be a meaningful and productive behavior. As early as 2003, Wyatt explicitly urged scholars to "take non-users and former users seriously as relevant social groups. . . who might influence the shape of the world" [**Wyatt, Oudshoorn, and Pinch2003**]. Moreover, in 2009, Satchell and Dourish similarly called for HCI researchers to consider non-users, and sought to dispel the notion that non-use is an "absence" or "negative space" [**Satchell and Dourish2009**]. A key theme in this literature is the relationship of the phenomenon of non-use to structural inequality across demographic groups [**Hargittai2007, Johnson et al.2016a, Redmiles, Kross, and Mazurek2016, Shaw and Hargittai2018**], a relationship we consider below.

The protest behaviors we study can be seen as a subset of the broader non-use phenomena observed and theorized in prior work. In a recent publication, Baumer et al. specifically emphasized the need to study different types of motivations for non-use [**Baumer2015**]. This present study can be understood as addressing this call, with our work focused specifically on non-use in protest of the values or actions of a technology company. One important recent contribution of the non-use literature has been to problematize the term “user” and even “non-use”. Specifically, researchers have called for treating non-use as a “continually negotiated practice” [**Baumer et al.2015a**] which is not characterized by a binary distinction between users and non-users [**Baumer et al.2015c, Baumer2018a, Baumer2015**]. In this view, the complex spectrum of use and non-use includes a variety of behaviors, e.g. deactivating an account, considering deactivating an account, taking a break from a platform, creating fake accounts, and many other behaviors [**Baumer et al.2013, Baumer2018a**]. Baumer and others [**Baumer et al.2015a, Baumer2018a, Baumer et al.2015b**] have adopted the term ”non/use” to encompass the spectrum of use and non-use behaviors, with “non-use” reserved for behaviors very close to one end of the spectrum. Our study reflects the complexity highlighted by Baumer and colleagues: we consider both people who remain users of a technology but protest by altering their use behavior and people who are protesting by ceasing their use entirely. As such, following Baumer et al.’s guidance, for the remainder of the paper, we leverage the term “protest non/use” when referring to the spectrum of behaviors exhibited by our respondents who are protesting a technology company. We use the term “protest non-use” when specifically referring to people who reported entirely stopping use of a technology. As we have above, we leverage the term “protest user” to describe all users who have engaged in protest non/use, as all people in this class are or were users of a technology. A large body of research on non-use and non/use investigated these behaviors' association with structural inequalities on a variety of platforms (e.g. [**Hargittai2007, Johnson et al.2016a, Redmiles, Kross, and Mazurek2016, Shaw and Hargittai2018**]), and this line of work informed our analysis and thinking of the relationships between demographic factors and protest non/use. For instance, using a sample of U.S. households and focusing on Facebook, Baumer showed that age, gender, and income are predictive of various types of Facebook non/use [**Baumer2018a**]. Below, in our Results section, we compare our demographic findings with those from Baumer and reflect on the implications of our observed demographic trends in protest non/use. Past research has also identified how individual and social factors relate to behaviors on the non-use end of the non/use spectrum, providing helpful lenses for us to interpret our findings. Guha et al. discussed how the lack of agency and control on Facebook plays a role in users leaving Facebook [**Guha, Baumer, and Gay2018**]. Baumer et al. identified a number of individual and social factors that predict reversion after leaving Facebook, including the concerns about impression management and friends' reactions [**Baumer et al.2015c**]. Lampe et al. found social capital is a strong negative predictor of whether somebody will join Facebook at all [**Lampe, Vitak, and Ellison2013**]. Although we did not collect or analyze these types of individual or social factors, we interpret and discuss our findings in light of the context provided by these studies. Finally, studies on privacy-driven behaviors have identified several forms of non/use that can be seen as protest non/use, directly influencing our construction and understanding of protest non/use. As privacy concerns are a prevalent motivation for non/use [**Baumer et al.2013,**

Lampe, Vitak, and Ellison2013, Stieger et al.2013], prior work has shown technology users adopt a variety of obfuscating strategies in protest, e.g. providing fake personal information [**Guha and Wicker2015, Sannon, Bazarova, and Cosley2018**]. Additionally, Mathur and colleagues' work on browser-based blocking extensions revealed some people's overwhelming discomfort with online tracking as well as their corresponding blocking strategies (e.g. using anti-tracking and ad-blocking extensions) [**Mathur et al.2018a**]. Our study bolsters these findings and we discuss the implications of protest users for privacy research and vice versa.

Protests against Technology Companies

Many recent protests against technology companies, such as Amazon, Uber, and Facebook [**Granville2018, Greenfield, Frier, and Brody2018, Kasperkevic2018, Semuels2017**], are similar to traditional consumer boycotts: a group of people withholds engagement with a company to attempt to force the company to change some practices. As such, the large body of research on consumer boycotts (e.g. [**King and McDonnell2015, McDonnell, King, and Soule2015**]) can provide important context for our work. There has been some research on participation rates and outcomes of consumer boycotts. Based on a survey of the American consumers, more than 28Protesting behaviors in the technology domain can take on various forms corresponding to the different ways tech companies generate revenue. For instance, advertisements are a primary source of profit for some major tech companies (e.g. Google and Facebook) [**Lotz2019**], whereas companies in other sectors sell products and services directly to consumers. Thus, protesting behaviors in the tech domain include avoiding visiting the website of an ad-driven tech company (e.g. boycotts against Facebook), refusing to purchase goods or services from a company (e.g. boycotts against Uber and Amazon), or disrupting an ad-revenue generating platform (e.g. the 2015 Reddit blackout by sub-Reddit moderators [**Matias2016b**]). In addition to these protesting behaviors that attempt to reduce a company's ad revenue, a "data strike", i.e. a group of users withholding their "data labor", can also negatively impact many profitable intelligent technologies. We unpack this form of protest in detail in the immediately following section. Our research is also motivated by recent interest in "boycott-assisting technologies" [**Li et al.2018b**] such as Buycott [**Buy**] and Out of Site [**Li et al.2018b**] that aim to facilitate consumer boycotts offline and online, respectively. In particular, these technologies emphasize the collective nature of boycotts and inform boycott participants of recommended actions and their collective outcome. As we discuss below, our study provides concrete design implications for designers of boycott-assisting technologies to specifically support people protesting tech companies.

Data Labor

Recent work [**Overdorf et al.2018**] has identified that protests like those we consider here may be especially powerful compared to protests against non-tech companies, making understanding the prevalence and motivations of protest users all the more important. This research highlighted how, due to the reliance of most tech companies on intelligent technologies, users of these companies' products generally have two roles, each with its own source of power: users are consumers of services with "consumer power" and users are also data-generating "laborers" with "data labor power" [**Vincent, Hecht, and Sen2019**].

The latter role emerges from the critical dependence on user-generated data of many tech companies' intelligent technology-driven core services (e.g. recommender systems, search engines). Protest users exercise their consumer power when they stop or change their use of a technology and thereby reduce their contributions to sales and advertising revenue. Protest users exercise their data labor power when their stopped or changed use of a technology results in fewer products being rated, fewer pages being liked, and/or less implicit feedback being collected, thus damaging profitable recommender systems, search engines, and related intelligent technologies. These two roles and their corresponding sources of power make protest users particularly influential relative to traditional protests against non-tech companies, in which participants largely only have consumer power.

Methods

This paper reports findings from two web-based surveys conducted in 2017 and 2019. The first survey was designed to broadly explore the prevalence of and the reasons for protest non/use (protest use and protest non-use). Our second survey focused on five prominent technology companies and elicited in-depth responses about motivations, tactics, challenges, and roadblocks associated with protest non/use. Both surveys used nationally representative sampling by a third party, as is common in large-scale studies that have examined non-use and non/use (e.g. [**Baumer et al.2013, Baumer2018a, Guha, Baumer, and Gay2018**]). Below, we present details about our survey design, recruitment methods, and respondents.

Survey Design and Recruitment

Two authors designed the first survey in October 2017, and it was intended to be exploratory in nature. It was funded by a large non-profit organization at which these two authors are employed. Respondent recruiting was completed by the professional survey company SurveyMonkey, which used its proprietary approach to generate a nationally representative sample of Internet users who live in the United States and were at least 18 years old. The survey was completed by 463 people and contained both fixed-response and free-response questions. The fixed-responses questions were generally targeted at understanding the prevalence and motivations of protest users, and the free-response questions were open-ended. Some of the demographic information about respondents in this survey came from SurveyMonkey, and the survey asked directly about respondents' political views. The results from the first survey indicated that a non-trivial portion of the public was engaged in protest non/use against tech companies (as reported below, 9% of respondents reported themselves as protest users of at least one prominent tech company). These results – along with increasing media coverage and public interest in protesting tech companies – motivated us to launch a second, in-depth, and more focused survey in 2019. All authors were involved in the design of the second survey. Building off the basic structure of the first survey, the second survey sought to acquire more detailed and structured information about protest non/use, as well as to update the top-line numbers to assess whether the ranks of protest users were growing. More specifically, our second survey was designed around two structured research questions:

RQ1 – Basic Descriptive Information: (a) What is the prevalence of protest non/use? (b) What are the motivations behind protest non/use? (c) What tactics are employed?

RQ2 – Challenges and Roadblocks: What challenges do protest users face and what roadblocks prevent people from becoming protest users?

Our 2017 and 2019 surveys have important differences, both in terms of the questions we asked and how the questions were specifically framed. We made a number of additions to the 2019 survey to obtain data more explicitly targeted at our research questions. In order to gather data to directly answer RQ1(c) and RQ2, we added questions about protest tactics, challenges protest users faced, and roadblocks faced by potential protest users. The answer choices for questions regarding challenges and roadblocks were drawn from the free-response answers provided by respondents to the first survey, as well as themes in the non/use literature and in media coverage of protests. Additionally, whereas the 2017 survey focused on multiple-choice questions with single answers, the 2019 survey was primarily based around multiple-choice multiple-answer (i.e. select-all-that-apply) questions with an option to provide free-text input to explain or expand upon one's answer. The 2019 survey also integrated answer choices that were not included in our first survey but were reported by 2017 respondents in the free-response questions (e.g. "the company profits from my data" and "I have concerns about the company's bias against gender, race, or other demographics" as motivations for protest non/use). Additionally, the 2019 survey included a Likert-type question about how difficult it is to protest a given company (on a scale from 1 to 5) after a respondent reported being a protest user of the company. In terms of how we framed the survey questions, although both surveys used the term "boycott" as a shorthand for "protest use and protest non-use" as we hypothesized this term would be much easier to understand for respondents, we altered the exact definition of "boycott" provided in the 2017 survey for the 2019 survey. In the 2017 survey, boycotting was defined for respondents as "deciding to stop using, or use much less of, a technology or company as a protest or statement, or because you disagree with the company's values." In 2019, we updated this definition to be "stopping or changing your use of a company's products or services, because you disapprove of the company's values or actions." This updated definition was meant to capture the many forms protest non/use can take against tech companies (as discussed in Related Work). As mentioned above, the 2019 survey asked additional questions about specific tactics compared with the 2017 survey, and some of these answer choices about tactics can be employed for non-protest reasons (e.g. private browsing and ad blockers might be used for reasons unrelated to protesting a company's values or actions). As such, we took care in survey design to avoid confounds surrounding the reasons for the use of a potential protest tactics. Respondents were first asked if they were protest users of a given company, and then they were asked which tactics they used in their protest. While this avoided confounds in our top-line numbers about participation rates, we did still see some confusion when respondents were enumerating the tactics that they used to implement a specific protest, and we discuss this more below.

2019 survey was conducted through Qualtrics (following prior research on non-use, see [8]), which also uses proprietary methods to perform nationally representative sampling (we detail the demographics of our respondents in Table 1). The survey was deployed in early 2019 by a subset of the authors who are employed at an academic institution, in accordance with their institution's IRB. This survey had 429 responses in total. However, we found that some responses appeared to be low-quality (e.g. free response fields filled with random characters). The first two authors examined all the responses independently to identify low-quality responses and then compared and discussed their findings to build a merged set of low-quality responses. In total, 31 responses were flagged as low-quality and were removed from all analyses, leaving us with 398 valid responses. Given that we modified the survey design and used two different companies for proprietary sampling, we must interpret any observed trend in the two survey results with some caution. However, considering that some differences between the two nationally representative samples' results are very large (e.g. the increases in our top-line participation rates), they very likely represented movements in the underlying phenomenon.

Respondent Demographics

Table 1 shows the demographic data we obtained from our surveys. In the 2019 survey, all demographic questions were optional, but 90% of the 2019 respondents answered all the demographic questions. Comparing Table 1's "All respondents" column with U.S. Census Bureau data [**Bureaub**], we find that our samples were reasonably balanced across a number of demographic factors, with a slight over-representation of the low- to middle-income population. The 2017 sample also has a relatively large share of respondents who are at least 60 years-old compared with the U.S. Census Bureau's population estimates (with 28% of the U.S. adult population being at least 60 years old) [**Bureaua**]. On the other hand, the 2019 survey has a relatively small share of this population.

Below, we constructed logistic linear regression models to further examine the relationships between protest non/use and demographics. Age and income were represented as ordinal variables using the levels shown in Table 1. Political stance and gender were represented as categorical variables.

	2017 Survey		2019 Survey	
	All respondents	At least one company	All respondents	At least one company
Age				
18 - 29 years old	20%	19%	26%	32%
30 - 44 years old	21%	22%	41%	42%
45 - 59 years old	25%	27%	23%	20%
60+ years old	35%	32%	10%	5%
Gender				
Female	55%	45%	48%	35%
Male	45%	54%	51%	63%
Non-binary	-	-	0.5%	2%
Agender	-	-	0.3%	0%
Transgender	-	-	0.3%	0%
Political stance				
Democrat	36%	30%	38%	43%
Republican	21%	14%	24%	22%
Independent	33%	38%	31%	31%
Other	10%	19%	6%	3%
Income				
< 25,000	18%	9%	23%	18%
25,000–49,999	25%	19%	33%	33%
50,000–74,999	19%	31%	20%	24%
75,000–99,999	12%	9%	10%	13%
100,000–124,999	13%	22%	7%	7%
125,000–149,999	5%	3%	4%	4%
150,000+	8%	6%	2%	1%

Table 1. Self-reported demographic information of respondents, broken down by the percentage of total respondents ("All respondents" column) and the percentage of respondents who were protest users for at least one company ("At least one company" column)

Respondent demographics.

Margin of Error and Confidence Intervals for Percentages and Instances

Using margin of error calculations for a random sample, each survey had a large enough sample to achieve a margin of error of 5% at a confidence level of 95% for our target population (web users in the United States who are at least 18 years old). Many of our results are simple percentages of respondents, such as the percent of users protesting a given company. For these percentages, following recent suggestions for reporting results in HCI research [**Dragicevic2016**], we compute non-parametric 95% confidence intervals (CI) using empirical bootstrap resampling (a popular approach for generating CIs for survey results [**Shao2003**]). Specifically, we used software from Beecher et al. [**Beecher et al.2019**] and used 10,000 resampling iterations for each CI.

Not all of our results are reported as percentages. For results relating to motivations, tactics, challenges, and roadblocks, our survey provided the numbers of instances of each motivation, tactic, challenge, and roadblock. An instance refers to one respondent reporting one motivation (or tactic, challenge, or roadblock) for one company. Thus, one respondent can have multiple instances spread across multiple companies. For example, one person might protest Facebook because of privacy concerns and the company's political stance, which would correspond to two difference instances of motivations (privacy and political). In our results, we report both the number of instances for each company, and instances summed across companies. These summed instances do not represent estimates about the national population, but instead represent how frequently a motivation, tactic, challenge, or roadblock was reported by our respondents, allowing for a single individual to contribute many instances. For these results, instead of reporting percentages with confidence intervals, we report only the total count of instances and interpret our results accordingly.

Results

Below we unpack the results from the surveys. As our 2019 survey was targeted specifically at our research questions, we focus primarily on our 2019 results below and provide the 2017 results for context. We first give an overview of the percentages of people who reported being protest users, and then detail the percentages of protest users for each company. We further unpack the motivations, tactics, challenges, and roadblocks associated with protest non/use.

Prevalence of Protest Users

The highest-level result from our 2019 survey is that a substantial share of respondents – 30% (CI: 25 – 34%) – reported being protest users. The majority of the protest users (21%) were protesting one company only, followed by 5% reporting two companies. Very few protest users were protesting more than two companies. 33% (CI: 28-37%) of respondents expressed interest in becoming protest users of at least one tech company against which they were not currently engaging in protest non/use, approximately half of whom (19% of respondents; CI: 15-22%) were not currently protest users of any company. In total, 48% of respondents (CI: 44-54%) were either actively engaging in protest non/use (30%) or were only interested in doing so (19%), after rounding to the nearest percent. Notably, the prevalence of active protest users we observed (30%) is very close to estimates of the prevalence of political consumption (i.e. boycotts and buycotts) in the United States in 2011 and 2012 (28%) [**Koos2012, Newman and Bartels2011**].

Figure 1 unpacks our results about the prevalence of protest non/use on a company-by-company basis. Facebook stands out as a particularly common target of protest users and potential protest users: nearly one-third of respondents reported that they were currently a protest user of Facebook or were interested in becoming one. In Baumer's 2018 study, 17.6% of respondents stopped using Facebook (through account deactivation) and 22.4% considered doing so, meaning 40% of respondents were, or considered, stopping Facebook use. Our observed number of active and potential Facebook protest users is thus slightly lower than Baumer's 2018 result. Note that in Baumer's study, the number of active and potential Facebook non-users included those who might not be protesting Facebook. Such respondents in our study would not identify themselves as protest users, potentially explaining our lower percentage.

Also of note in Figure 1 is that Amazon, Google, and Microsoft have more potential protest users than actual protest users, suggesting a lower protest "conversion rate" for these companies. Below, we present the specific roadblocks reported by potential protest users of these companies. These roadblocks may play a role in influencing the conversion rate of potential protest users to active protest users. Table 2 puts our top-line results from the 2019 survey in context with those from 2017. Whereas 30% (CI: 25-34%) of respondents in 2019 reported being protest users of at least one company, the equivalent number in 2017 was only 9% (CI: 6–11%). In particular, we see significant increases in protest rates of Facebook and Apple. The percentage of respondents protesting Facebook more than tripled in 2019 from 5% to 18%, and the percentage for Apple in 2019 is four times that of 2017, going from 3% to 12%. The remaining three companies, Microsoft, Amazon, and Google also see an increased rate of protest users, with the percentages roughly doubling. Overall, we see rising protest rates across all five companies, but Facebook and Apple see the largest increases. Recall that these comparisons need to be interpreted with caution: the two surveys were not identical in design or sampling (see Methods). Furthermore, differences in protest prevalence rate will be affected by changes in company user bases (e.g. people who didn't use Facebook at all in 2017 may have joined Facebook and engaged in protest use in 2019). Nonetheless, the size of the delta we observed suggests that the prevalence of protest non/use has increased in the last two years.

Who Are Protest Users?

According to our 2019 data, certain groups are more likely to protest: it appears that respondents who identified as male protested more than other gender identities, and younger respondents protested more than older respondents. A logistic regression that uses self-reported demographics as the independent variables and protest non/use for at least one company as the dependent variable suggests that both of these are statistically significant associations (p /¡ 0.05, see Table 3). In particular, male respondents were 2.4 times more likely than female respondents to protest when holding other factors constant, which is consistent with the descriptive statistics in Table 1. With one increment of the age groups in Table 1, older respondents were only 0.632 as likely as younger respondents to protest.

YEAR	FACEBOOK	APPLE	MICROSOFT	GOOGLE	AMAZON	TOTAL
2017	5%	3%	2%	2%	2%	9%
2019	18%	12%	6%	5%	6%	30%

Table 2. The percent of protest users against five major tech companies in 2017 and 2019.

The percent of protest users against five major tech companies in 2017 and 2019.

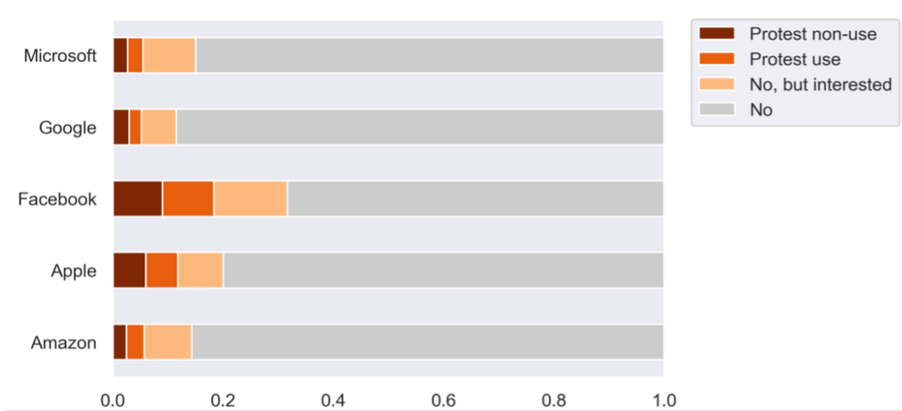


Figure 2. Protest non/use against five major tech companies from our 2019 survey. The x-axis indicates the fraction of respondents who engaged in protest non-use, protest use, or were interested in becoming protest users. Respondent participant rate.

	Coef.	*Std. err*	*p*	*Odds Ratio*
Intercept	-0.237	0.351	0.499	0.789
Political stance - Independent	-0.115	0.283	0.683	0.891
Political stance - Other	-2.017	1.059	0.057	0.133
Political stance - Republican	-0.066	0.318	0.836	0.936
age	-0.459	0.141	0.001 *	0.632
income	-0.035	0.076	0.649	0.966
gender - male	0.892	0.258	0.001 *	2.439

* indicates p-value less than 0.05.

Coefficients and odds ratios for a logistic regression with self-reported demographic information as independent variables and engagement in protest non/use of any company as the dependent variable. The pseudo R-squared of the model is 0.06.
Table 3. Who are protest users overall?

With respect to Facebook specifically, analogous to Baumer's finding that younger respondents are more likely to deactivate their Facebook account [2], our model (Table

	Coef.	Std. err	p	Odds Ratio
Intercept	-0.786	0.412	0.056	0.455
Political stance - Independent	-0.752	0.349	0.031 *	0.472
Political stance - Other	-27.891	476144.631	1	0.000
Political stance - Republican	-0.360	0.370	0.330	0.698
age	-0.390	0.168	0.020 *	0.677
income	-0.059	0.091	0.516	0.943
gender - male	0.901	0.311	0.004 *	2.462

Table 4. Who are Facebook protest users?

Analogous to Table 3, but with protest non/use of Facebook specifically as the dependent variable. The pseudo R-squared of the model is 0.07.

4) shows that younger respondents are more likely to be protest users of Facebook than older respondents. However, in contrast with the insignificant relationship between gender and Facebook deactivation that Baumer observed, men in our study were 2.4 times more likely than women to protest Facebook (very slightly more than our result for overall protest users). This difference may be due to the divergence in the definitions of protest users and non-users as mentioned above. In other words, although men and women are equally likely to be Facebook non-users, men may be more likely to do so as an action of protest than women. Also of note is that compared with Democratic respondents (the default intercept in Table 4), Independent respondents were less likely to protest Facebook (odds ratio=0.472), a relationship that we do not observe in the model considering all companies.

Motivations for Protest Non/Use

Our 2019 data provides us with rich information about motivations for protest non/use, with active and potential protest users selecting two motivations per company on

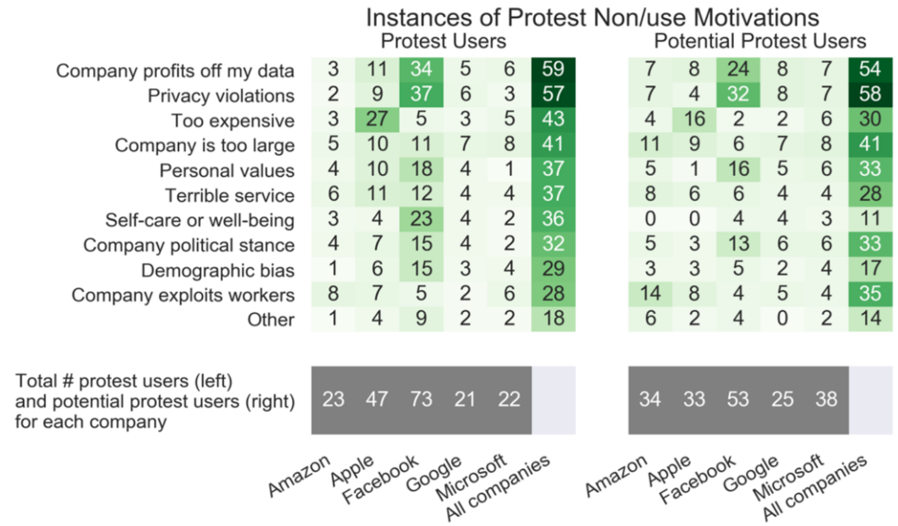


Figure 3. Instances of protest non/use motivations reported by our respondents who were protest users ("Protest Users", left) or interested in protest non/use ("Potential Protest Users", right). Each respondent could select multiple motivations. Includes the total number of (potential) protest users per company in grey for context.

Motivations

average (respondents could select all motivations that applied). We focus here on reporting the number of instances of each motivation, where an instance is a single motivation for protesting a single company (selected by a single respondent).

Figure 3 shows our motivation findings in detail. The left side of Figure 2 shows our motivation-related results for active protest users. The right side shows the equivalent findings for potential protest users, i.e. people who expressed interest in becoming protest users of a given company but were not doing so currently. Examining the left side of the figure, we see that the most-common motivations for actively protesting were concerns around companies profiting off of user data (59 instances) and privacy (57 instances). In other words, respondents indicated 59 times that they were motivated to actively protest a tech company because it was profiting off of user data and did the same for privacy 57 times. The next two most common motivations were cost (43 instances) and company size (41 instances). The most prominent motivation for protesting, concern about companies profiting off of user data, does not align with prior work which has suggested that college students cared little about how their data is used by platforms [**Young and Quan-Haase2013**] and placed very small monetary value on protecting data [**Grossklags and Acquisti2007**]. One reason for this difference may be the increasing awareness of data-driven business models in the past few years. The qualitative data from the 2017 survey was an early signal that profiting off of user data might be a prominent motivation. Some respondents from 2017 took strong stances on the topic, saying "I resent the invasive tentacles of tech companies. They are trying to control and profit from everything we do in life. They don't respect privacy they just want $$" and "they sell my personal information exploiting ME MAKING PROFIT OFF OF ME, without giving me any financial share of their profit pirating." Our quantitative data from 2019 suggests that these sentiments are spreading more broadly. The prevalence of privacy concerns visible in Figure 3 resonates with HCI studies of privacy and surveillance (e.g. [**Sannon, Bazarova, and Cosley2018, Young and Quan-Haase2013**]). In particular, Baumer et al. found in 2013 that the top motivation for leaving Facebook or limiting Facebook use was privacy. Our results suggest that, six years later, these concerns remain serious for people who engage in various types of protest non/use of Facebook (including leaving Facebook). Indeed, examining Figure 3, we see that of the 73 users who reported being active protest users of Facebook, 37 (51%) indicated that they were doing so for privacy reasons. We see a similar trend on the right side of Figure 3, where privacy was the number one motivation for being interested in becoming a protest user of Facebook

(60% of potential Facebook protest users). The reported motivations in the two surveys have some other differences, although we did not provide identical options and therefore direct comparisons must be interpreted with substantial caution. For the options that overlap between two surveys, privacy concerns remained the top motivation in aggregate. However, the second-most-popular overall option in 2017, disagreeing with the company's political stance, substantially diminished in prominence in our 2019 data. Furthermore, looking at these trends per company, we observe a large increase in people protesting Amazon because of working conditions, perhaps relating to the media's coverage of the issue (e.g. [**Kasperkevic2018, Pri, Why**]).

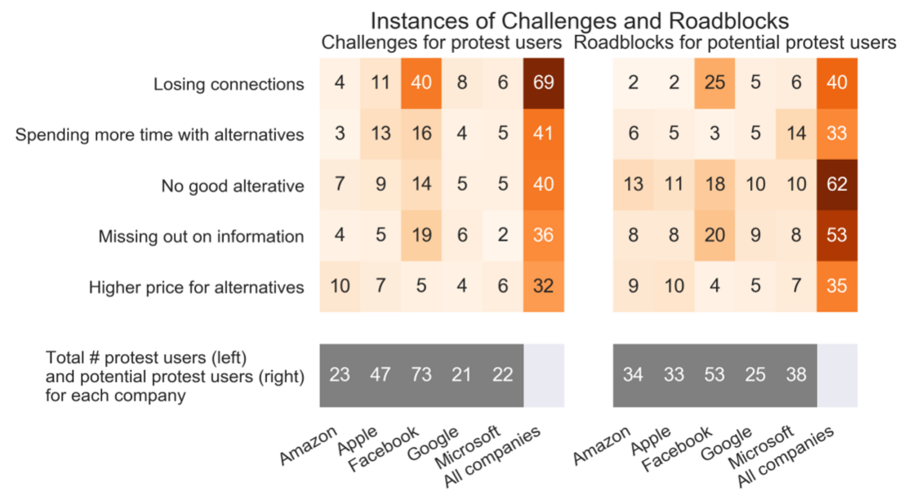
The Tactics of Protest Users

Our 2019 survey elicited information on the specific tactics leveraged by protest users in their protest non/use. Overall, non-use was the most-common reported tactic. 93 instances of non-use were reported in total, where an instance in this case means that a single respondent reported entirely halting the use of a single company's products. Respondents also reported 129 specific instances of protest use overall, i.e. still using a technology but with protest tactics, including ad blocking, private browsing, using fake accounts or fake data, using anti-tracking extensions, and using products while logged out. Among these protest use instances, we observed that using ad blocking (41 instances) was the most common tactic. The prevalence of ad blocking is not surprising given a recent survey on Amazon Mechanical Turk (MTurk) showed over half of participants use ad blockers [39]. Also consistent with the MTurk survey, the use of anti-tracking extensions was less prevalent than ad blocking in our study, with 18 instances of anti-tracking reported. Following using ad blockers, providing fake accounts or data (27 instances) and using private browsing features (24 instances) were the second- and third-most prevalent tactics among our respondents. These tactics largely overlap with privacy-driven obfuscation approaches that have been reported in privacy and surveillance research. For instance, Sannon et al. found that 21.9% of their recruited respondents lie to computing systems to protect their privacy [49]. Our results suggest that protest users were re-appropriating these privacy-protection strategies as a means of protesting, indicating an overlap in tactics among protest use and privacy protection. This overlap may have important implications that we unpack in Discussion. Focusing on tactics related to protest non/use of Facebook specifically, similar to prior work showing the non-binary nature of Facebook non/use [5], our survey responses imply that protesting Facebook involves nuanced behaviors that are not limited to simply deleting or deactivating one's Facebook account. Among the 37 respondents who were using Facebook but engaging in protest use, using ad blockers (16 instances) was the most common tactic, followed by using anti-tracking extensions (11 instances) and private browsing (9 instances). In our survey, 53 respondents who reported protest non-use ("stopping entirely") of a specific technology also selected additional protest tactics against the company, e.g. using ad blockers and private browsing. This may indicate very nuanced tactical strategy (e.g. people who stop using Facebook and also use private browsing or anti-tracking to attempt to avoid Facebook tracking on other websites) but might also indicate confusion on behalf of a respondent (e.g. perhaps people who used an ad blocker for reasons unrelated to protest of a specific company were confused by this question). As our data did not fully explain this behavior, our reported results include only the protest use tactics used by people who indicated that they continued to use a technology.

Challenges and Roadblocks

Figure 4 presents the challenges reported by active protest users of each company on the left and the roadblocks reported by potential protest users of each company on the right. Unlike was the case for protest non/use motivations, there was a notable difference in responses between those who were actually protesting and those who were interested, but not doing so. Here, we see concerns about "losing connections" was by far the most common challenge for active protestors (driven by people protesting Facebook). On the other hand, and raising important concerns related to the discussions around the possible monopoly power of some technology companies, the lack of alternative products was the most common roadblock to protest non/use for interested respondents (across all the companies).

Missing out on information and losing connections, the two major challenges reported by protest users of Facebook, are consistent with prior work [5]. As Facebook is primarily a social networking site, it is unsurprising that these two options, which represent social challenges (as opposed to economic or technical challenges) are common among active and potential protest users. In the case of Amazon, we see that paying higher prices for alternatives was the top challenge for active protest users, but respondents who were interested in protesting Amazon identified the lack of alternatives as the top roadblock.



Challenges active protest users reported (left) and roadblocks that potential protest users reported (right). Includes the total number of (potential) protest users per company in grey for context.

Figure 4. Challenges

This disparity suggests that (perceptions of) higher prices may be a roadblock for some, but a manageable challenge for others, hinting at a role of socioeconomic status in the ability to become a protest user. We discuss these results further below, putting them in the context of related findings from other studies of non/use and non-use (e.g. [59]). Amazon was rated as the most difficult to protest by active protest users, with an average difficulty of 2.4 on a 1 to 5 scale (with 1 corresponding to “very easy” and 5 to “very difficult”) and Apple was rated as the least difficult, with an average of 1.7. For other companies, the average rating was around 2 (“easy”) or lower. Overall, it seems our active protest users did not find it to be especially difficult, although data beyond a single Likert-type response will be important to confirm this result.

Discussion

At a high level, our survey results suggest that protest users have become a substantial force in the sociotechnical landscape. Although our 2019 survey was small and is just one survey, we observed that three of out of every ten respondents are already protest users, and another almost one-fifth of the respondents have an interest in becoming protest users. These results – along with the more detailed findings about motivations, tactics, challenges, and roadblocks – have important implications for a variety of stakeholders, including researchers in social computing and other areas of computing, technology designers, and institutions that own prominent technologies. We discuss some of these implications below.

Technologies to Support Protest Users

As noted above, the social computing literature and wider computing community have become increasingly interested in developing technologies to support protest non/use (e.g. boycott-assisting technologies [**Buy, Li et al.2018b**], protective optimization technologies [**Overdorf et al.2018**]). One of the most significant implications of our results is that they suggest that there is a truly substantial "market" for these technologies. Our findings indicate that this market may include up to almost half of American Internet-using adults, providing substantial support for more research and development in this area. Additionally, our findings also present something of a partial roadmap for new technologies to support effective protest non/use. For instance, our results highlight the importance of future technologies that can offload the burden of finding and using alternative products for protest users and thereby lower the threshold to participate in protest non/use. Such tools may meaningfully increase the percentage of people who can actualize their desire to become protest users against the target (i.e. move from the right side to the left side of Figure 3). Although these tools may adopt a number of different approaches, one approach might be to act as an intermediary to a desired service (e.g. purchasing some product), directing people to alternatives whenever possible. For example, a browser extension could autonomously route shopping queries away from a targeted company, with that targeted company being a backstop if there truly is no other company offering the product at a similar price. One could also imagine a similar tool for web search that routes search queries to minority players like DuckDuckGo when those queries reflect information needs that are straightforward to satisfy (e.g. navigational queries like "CSCW 2019"). The large number of existing protest users amongst our respondents and the wide variety of tactics employed also introduce a promising opportunity for "computer-supported collective action" [**Shaw et al.2014**]. For instance, new tools could help to identify and mobilize protest users who have the most leverage over the target (e.g. influential members of a social network, people who contributed especially valuable data, etc.). These tools could also make suggestions to existing protest users about particular days to avoid a platform (i.e. a day-long boycott) or specific types of fake data to provide. Additionally, the prevalence of ad blockers and anti-tracking extensions among protest users suggests that these tools could also coordinate collective action to make individual protests more effective. In particular, as visible progress of collective action sustains participation [**Ling et al.2005, Shaw et al.2014**], current ad blockers and anti-tracking browser extensions may consider communicating how many protest users are taking action and estimates of the protest's impact on web traffic [**Li et al.2018b**] or ad revenue [**González Cabañas, Cuevas, and Cuevas2017a**] (e.g. "Over the last week, n other people have also been using this anti-Facebook tracking extension and m ads have been blocked, costing the company p dollars").

Designing and Studying with Protest Users in Mind

The results of our survey point to the need for researchers and developers to consider protest non/use in the technology design process. This would involve asking questions like: How and why might people contest a new feature or system? Are there ways to account for this contestation before it starts? How resilient would the system be to such contestation? Designing with protest users in mind may be a useful approach to shift designers' attention to how people might negatively react to technology and means building systems that recognize the value and power of all technology stakeholders, including users, protest users, and non-users. This is an approach that would further supplement existing user-centered design approaches, such as participatory design [**Muller2003a**] and value-sensitive design [**Friedman and Nissenbaum1996**], and relates to the notion of "heuristic preventive design" introduced at CSCW last year [**Li et al.2018b**]. On a related note, social computing researchers also need to be aware of protest users as a dimension (and potential confound) in studies of large-scale online platforms. For instance, our results suggest that a study of Facebook use in the United States may want to consider how the research questions and chosen methodologies (e.g. recruiting through Facebook ads) might be affected by protest non/use. More generally, as the growing literature around social media and other technologies emphasizes the demographic gap in technology use, future work should particularly account for the potential influences of protest non/use on this gap. For instance, will the demographics of Facebook users change because the younger population protest more?

Protest Users and Technology Non/Use

Our study unpacks the subset of technology non/use behaviors driven by protest, directly responding to Baumer 2018's call for examining "relationships between different form of technology non-use and different types of motivations" [**Baumer2018a**]. As is discussed above in more detail, our work also points to potentially unique characteristics of protest users with respect to (non)users who are considered in studies about more general non-use and non/use. For instance, although men and women are equally likely to deactivate Facebook, our results suggest that men are more likely to be protest users of Facebook. Similarly, privacy drives both protest users and (non)users to change their Facebook usage or leave Facebook, but protest users are uniquely concerned about Facebook profiting off of their data. More generally, while our paper maps out a new territory within the domain of non/use, our paper also highlights the need for more targeted research on the relationship between protest use, non/use, and non-use.

Protests, Privacy, and Surveillance

Viewed through the lens of the relevant privacy literature (e.g. [**Guha and Wicker2015, Masood et al.2018, Mathur et al.2018a, Sannon, Bazarova, and Cosley2018**]), our findings point to an interesting overlap between protest use and privacy-driven behavior, an overlap that is fertile ground for future empirical and theoretical work. In particular, the exact same tactic – e.g. using fake accounts / data and private browsing – can be deployed either as a means to protect individual privacy or as a means to protest a company that makes money off of personal information or data labor [**Mathur et al.2018a**]. Likely, in many cases the tactic is the result of both motivations at the same time. This overlap highlights that actions that protect one's privacy may go beyond simple self-interest and are affected by complex sociotechnical contexts, e.g. the company's business model and public image. It also suggests the reverse: the literature on protests against technology company has been dominated by a collective action frame, but there may also be highly self-interested benefits and motivations to these protests. The overlap between protest non/use and privacy-driven behaviors may additionally present promising opportunities to leverage existing privacy protection tools for protesting purposes. For example, AdNauseam, a browser extension that simulates random clicks on ads to obfuscate tracking by online advertisers, may facilitate protests against technology companies by automatically generating fake data to create "garbage" inputs to trained models [**Howe and Nissenbaum2017a**]. Future work might seek to estimate the economic and social effects of widespread obfuscation-based protests. Additionally, the reported privacy-driven behaviors by protest users to avoid tracking by tech companies suggest that future work may also want to examine protest non/use through a lens informed by theories of surveillance [**Albrechtslund2008, Guha and Birnholtz2013**]. In particular, past work from Albrechtslund has contrasted vertical "Panopticon / Big Brother" concepts of surveillance (in which there exists a hierarchy of "watchers" and "watched") with horizontal "participatory surveillance" [**Albrechtslund2008**]. The participatory surveillance framing argues peer-to-peer surveillance by social networking users is a form of maintaining friendship and thereby empowering, playful, and positive

[**Albrechtslund2008**]. These potentially conflicting approaches to conceptualizing surveillance suggest conceptual complications faced by protest users. Protest users' obfuscating tactics (e.g. fake data, fake accounts) to resist vertical surveillance may hinder their participation in social surveillance, as they withhold data from target technologies. This is particularly interesting when considering protests against social network companies like Facebook. For example, the Facebook protest users who reported providing fake data to Facebook in our study may not see certain content with which their friends have engaged and thus lose the opportunity to participate in the positive aspects of social surveillance (while simultaneously receiving some protection from the negative aspects of vertical surveillance). Similarly, the Facebook protest users who reported entirely halting the use of Facebook (e.g. protest non-use) or contributing fake content (a protest use tactic we observed), may lack the opportunity to make connections with people that share similar interests. Future work should further investigate how protest non/use influences one's ability to engage with social surveillance.

Protest Users and Intelligent Technologies

Prior work on collective action campaigns suggests that protest users may be particularly effective at impacting intelligent technologies. Vincent et al.'s work identified two types of collective action campaigns that have the potential to meaningfully reduce the performance of highly-profitable intelligent technologies like recommender systems: "data boycotts" and "data strikes" [**Vincent, Hecht, and Sen2019**]. Both of these campaigns map closely to the phenomena studied here. Boycotts correspond directly to protest non-use. Some of the behaviors (e.g. anti-tracking) observed in our survey could be used to contribute to a data strike. Given the close correspondence of protest non/use, data strikes, and data boycotts, the observed prevalence of protest non/use should be of significant concern to companies that use data-driven intelligent technologies. According to Vincent et al.'s research, boycotts and strikes in which 30% of the user base participates - the prevalence of protest users that we observed – can meaningfully reduce the performance of a recommender system for the 70% of the user base that does not protest [**Vincent, Hecht, and Sen2019**]. As such, given their prevalence, protest users are already likely reducing the performance of intelligent technologies owned by targeted companies, even for people who are not protest users. If the scale of protest non/use grows, Vincent et al.'s work suggests that this effect will continue to increase.

Protest Users and Monopolies

A concerning result in our survey is that many people felt they could not stop or change their use of a given technology because there were no alternatives to this technology. This finding provides a data point for the growing discussions about monopoly power of many of the companies in the technology industry [**Herndon2019, Manjoo2018, Martínez2019, Wright2018**]. If a user of a technology cannot "put their money where their mouth is" due to the lack of competitors, this supports an argument that there has been a market failure. It may be that much of the protest use we observed would become protest non-use if there were more competitors available. Indeed, this is the motivation for Vincent et al.'s "data strike" concept: data strikes allow people to continue to use a platform while exerting some leverage over it. Overall, it is clear that more research is needed on the relationship between protest use, protest non-use and market competition.

Our results provide a useful data point on this relationship, but they come from just one survey of limited size and scope.

Future Work on Protest Users

At the most basic level, our findings highlight the need for follow-up work that examines the prevalence and character of protest non/use in more detail. This would involve in-depth qualitative research with protest users (and potential protest users), examining protest non/use in more diverse geographic contexts (see Limitations below), and even perhaps running larger-scale surveys. Following prior work on non-use (e.g. [**Baumer et al.2015a, Baumer2018a, Hargittai2007**]), social computing research should also examine protest non/use explicitly through a socioeconomic lens. Our results suggest that there are complex socioeconomic contours associated with protest non/use. In particular, there are hints in our results of protest non/use being a privilege of people who can afford it, with lack of alternatives being the most common roadblock to catalyzing interest in a protest into action. In the terms of Wyatt's distinction between voluntary and non-voluntary non-use [**Wyatt, Oudshoorn, and Pinch2003**], our study reinforces that technology use can be non-voluntary as well. That is, our study provides early evidence showing potential protest users were "stuck" using technologies that they were interested in protesting. These results call out for future work to further investigate the role of socioeconomic factors in protest non/use.

Limitations

A major limitation of our study was that we sampled only adult web users in the United States. Of course, this population's protest non/use is of interest to many stakeholders: this population is both large in absolute number and is an important revenue source for prominent tech companies [**Flynn2018**]. However, we observed – as have others (e.g. [**Hargittai2007**]) – that technology non/use behavior varies with respect to demographics, prominent tech companies vary around the world, and our population is a small portion of overall web users. Future work should investigate how the prevalence, motivations, and tactics of protest users change across the globe. The challenges and roadblocks facing protest users and potential protest users will likely also be another source of important geographic variation. Although our use of third-party services to collect nationally representative data was appropriate for our early-stage contribution to the discussion around protest non/use and is a standard practice in the social computing literature (e.g. [**Baumer2018a**]), this approach limits our ability to validate our results. Given that our major findings are based on descriptive results with large effect sizes, it seems unlikely that this is a major validity threat. However, any fine-grained results from our surveys or similar surveys must be taken with a grain of salt and precise estimation about specific phenomena (e.g. "how many people use Private Browsing to view Facebook pages?") are likely inappropriate given the nature of our data. Our major findings relied on multiple-choice multiple-answer responses. Although we aimed to cover a wide variety of possible answers motivated by themes in the news media, the literature, and our 2017 free text responses, it is possible we missed certain answers or worded them in a way that confused respondents. We mitigated this risk through the inclusion of an "Other" option in most questions and did not see evidence of major omissions in those responses. That said, we must assume there is some risk of design error on top of any sampling error. Finally, it should be reiterated that the design differences in our two surveys provide important context for any comparison between the 2017 results and the 2019 results. We adjusted our survey design for the 2019 survey to more directly answer our research questions about protest users' motivations and challenges instead of deploying an identical survey. We also used two different survey companies, each with its own proprietary sampling approach. As noted above, these decisions led to us placing more emphasis on the descriptive statistics from the 2019 survey than on any direct comparisons between the two surveys.

Conclusion

In this paper, we present the results of two surveys that explore if, how, and why people stop or change their usage of major technology companies' products as a form of protest (we call such people protest users). We find evidence that such behavior is increasingly common (almost half of our respondents were protest users or interested in becoming protest users), and driven by a variety of motivations, particularly concerns about privacy and business models that profit from user data. Moreover, our survey highlights common tactics that protest users employed in protest, and the challenges and roadblocks that inhibited these protests. This work provides important context for the growing discussion around the relationship and power dynamics between the public and technology companies. We present design implications for new technologies to better support protest users and highlight important follow-up social computing research into their protesting behaviors.

Chapter II: Data Labor - A Taxonomy

Data generated by the public (e.g. behavior logs, user-generated content, and personal information) is primarily governed by just a small set of large, for-profit technology companies that consequently reap the bulk of its benefits (e.g. insights, predictions, and ad sales). Members of the public generate large troves of data in their daily interactions with technologies, e.g. behavioral logs, user-generated content, and personal information. Currently, this data is primarily stored by just a small set of large, for-profit technology companies that reap the bulk of its benefits (e.g. insights, predictions, and ad sales). Those who produce data for the technology industry have little to no power in deciding how their data is used or who it benefits [**Arrieta-Ibarra et al.2018, Lanier and Weyl2018b, Posner and Weyl2019**]. The power imbalance between the public and technology operators has manifested in public outcries about various industry practices in the tech sector. For example, users have little to no power to change corporate surveillance practices and monetization of user data [**Li et al.2019b**]. Similarly, those that produce valuable content such as open-source code and Wikipedia articles have no way to control how the fruits of their labor are being repurposed by developers of machine learning models like Copilot and large language models such as GPT-3 [**Brown et al.2020b**].

Given the public's lack of power over the data it generates, researchers, policymakers, and activists have advocated for rethinking elements of the data economy so the data-generating public can have a much stronger voice in the use and governance of data [**Viljoen2020, Arrieta-Ibarra et al.2018, Lanier and Weyl2018b**]. In particular, understanding data generation as a form of labor, or "data labor"[**Arrieta-Ibarra et al.2018**] is gaining traction as one potential approach to achieve this goal. Supporters of this approach have argued that the conceptualization of data generation labor will help the public to actively leverage their role to influence technology companies [**Vincent et al.2021a**]. This proposal has also generated discussion about how the recognition of data labor may play out over time. Proposals include supporting "data unions" [**Posner and Weyl2019**] or "mediators of individual data" [**Lanier and Weyl2018b**] that negotiate data use terms with technology firms on behalf of their data-producing "union" members [**Posner and Weyl2019**], drafting legislation that would grant users greater control over the data they produce [**Por2018, the**], and creating tools to support user-driven collective action [**Das et al.2021, Vincent et al.2021a**].

Despite these abstract "blueprints" for reconceptualizing data generation as labor, the research and policy community is missing the transformation of these blueprints into more concrete and actionable guidelines. Given the myriad of ways that data producers contribute to data-driven technologies, a clear characterization of data labor is crucial to guide researchers, data producers, and policymakers to realize their goal: addressing the power imbalance between the public and large technology firms in the use and governance of data. More specifically, such characterization will accentuate how different types of data labor may require different approaches to empower through research, development, and policy practices.

This paper provides *an actionable road map* for researchers, activists, and policymakers to empower data producers to shape the use and governance of data by identifying and characterizing five key dimensions of data labor: *visibility , end-use awareness, collaboration requirement, openness, and replaceability.* We focus on data labor that benefits large, for-profit technology companies because of these companies' vast accumulation of capital from monetizing the public's data and their subsequent influence on social outcomes. Reviewing the rich HCI/CSCW scholarship on computer-mediated labor and the related concepts of digital labor, and crowd/gig work, we provide a definition and a characterization of data labor. While the prevalent interpretation of *labor* implies intentionality and compensation, our definition adopts the HCI/CSCW scholarship's construction of labor and encompasses both intentional, compensated data production (e.g. labeling images on Amazon Mechanical Turk for a computer vision company) and unintentional, uncompensated data production (exposing one's personal preferences while using commercial recommender systems). The directions our road map provides are informed by complementary frameworks of empowerment introduced by Schneider et al. [**Schneider et al.2018**] and data leverage from Vincent et al. [**Vincent et al.2021a**].

Our road map of data labor provides a framework for studying and empowering data producers in sociotechnical systems. Our definition and five dimensions of data labor do not aim to be all-encompassing; rather, they serve as an important step towards better conversations about data labor and towards redistributing the decision making power away from technology companies towards the data-generating public. As unpacked below, efforts to empower the data-generating public would benefit from HCI and CSCW scholars' knowledge and expertise in understanding, assessing, and organizing human labor in sociotechnical systems. More broadly, our work is in line with recent influential scholarly efforts in developing taxonomies and frameworks to guide technology research and development (e.g. [**Chancellor et al.2019a, Selbst et al.2019, Vincent, Hecht, and Sen2019, Quinn and Bederson2011**]). We discuss opportunities for future work to identify potential dimensions.

Definition and Related Work

Defining Data Labor

Discussions of "data labor" have not yet coalesced on a concrete definition and largely operate at a conceptual level. In 2018, Arrieta-Ibarra asked "should we treat data as labor?", in response to the lack of recognition of users' role in the advancement of technology and the data economy [**Arrieta-Ibarra et al.2018**]. Recently, this approach has motivated economists, legal scholars, and computing researchers to explore potential implications of treating user activities as labor via simulations [**Bergemann, Bonatti, and Gan2021**]. For example, Jones et al. simulated how granting users' rights to the data they produce and allowing the data to be used across firms can maximize social gains from the data economy [**Jones and Tonetti2020**]. Others have taken a step further and recommended establishing third-party intermediaries that are analogous to labor unions to facilitate the relations between subgroups of users and technology companies [**Lanier and Weyl2018b**]. In a similar vein, practitioners have piloted applications and platforms that allow users to control who has access to the data they generate (the Solid project[2] and Streamr[3]). The data labor notion has also been viewed as a potential means to support fairness and accountability outcomes: vincent2021data laid out pathways by which data producers can influence the performance of computing technologies and argued that all technology stakeholders need to recognize and account for data producers' role in the booming data economy [**Vincent et al.2021a**].

Building upon these conceptual proposals that explicitly examine data labor in the tech industry [**Arrieta-Ibarra et al.2018, Jones and Tonetti2020**], we offer a working definition of data labor:

Activities that lead to digital records useful for capital generation.

Said differently, an activity must meet two criteria to be labeled as data labor: 1) it creates or enhances data, and 2) the resulting data helps an organization generate capital. Correspondingly, those who contribute data labor are "data laborers".

In this dissertation, we primarily focusing on the data labor that supports for-profit technology companies, directly or indirectly. Governmental agencies (e.g. census bureaus), research organizations, civil societies, and non-profit organizations also use data labor to generate capital (see Discussion for details); however, the power inequity between data labor and technology operators is more prominent in the case of for-profit technology companies–the focus of our analysis. Moreover, data labor for for-profit technology companies is also the primary force contributing to capital generation.

With data playing a more prominent role in technological progress and the tech sector (e.g. [**Brown et al.2020b, Vincent, Johnson, and Hecht2018**]), we foresee that more and more data generation activities will fall under this definition of data labor. As

[2]https://solidproject.org/
[3]https://streamr.network/

we will discuss much more below, it is possible (and indeed very common) that under this definition, many people are performing "illegible data labor", in which it is not apparent to data producers that data labor is being performed, such as writing comments on social media.

Beyond Arrieta-Ibarra's work [**Arrieta-Ibarra et al.2018**], other areas of computing have explored complementary concepts to data labor and therefore informed the definition above. Below we relate our notion of data labor to similar concepts of labor and work in literature.

The Evolving Definition of Labor in HCI and CSCW

The fields of human-computer interaction (HCI) and computer-supported cooperative work (CSCW) have long histories of studying computer-mediated labor across domains, contexts, and culture. We examined the activities that HCI and CSCW scholars consider to be labor to inform our definition of data labor. In other words, by examining the rich tradition of studying labor in HCI and CSCW, can we draw clear boundaries that delineate "data labor"?

To better understand what activities scholars have considered as labor in HCI and CSCW, we conducted a literature review of full papers published in CHI and CSCW–two premier venues for HCI and CSCW–that study human labor in sociotechnical systems. Specifically, using the ACM Digital Library Search, we examined papers in the CHI and CSCW conferences and PACM-HCI between 1982 and 2022 that mention the word "labor" (or "labour") in their abstracts. Our intention is to capture examples of labor that are less obvious/explicit than conventional labor activities such as crowdwork and desk work. Our search results are by no means an exhaustive list of labor activities as defined by HCI and CSCW scholars. We manually screened abstracts to determine whether the authors refer to the human behaviors they studied as labor. We removed papers that do not explicitly view the studied activities as labor, e.g. merely mentioning "division of labor". In the end, our corpus includes 55 papers from CHI and 46 from CSCW (see Supplemental Material for details about these papers).

Table 5 provides an overview of the types of computer-mediated labor that frequently appear in our corpus, encompassing both compensated work and uncompensated activities and thus informing our definition of data labor. While compensated work is a common type of labor that researchers study, including deskwork, gig work (driving for Uber/Lyft), crowdwork (Amazon Mechanical Turk tasks), and low-wage work [**Dombrowski, Alvarado Garcia, and Despard2017**], the scope of labor started to include human activities that occur outside of workplaces or work contexts since the early 2000s and expanded to "settings in everyday life", ranging from domestic labor, to leisure activities, to social networking [**Crabtree, Rodden, and Benford2005**]. For example, To et al. studied how the act of looking for social support through communicative technology after experiencing racism is a form of emotional labor among minority students [**To et al.2020**]. Using Wikipedia as a case study, Geiger et al. highlighted the immense number of "labor hours" volunteer editors contributed to the world-wide encyclopedia [**Geiger and Halfaker2013a**]. In summary, although the scope of computer-mediated labor was once hotly debated [**Schmidt2011, Crabtree, Rodden, and Benford2005**],

Table 5. An overview of computer-mediated labor

Monetary compensation?	Sub-category
Yes	Desk work and office work
	Gig work and crowd work
	Low-wage work and service work
	Public service work
	Content creation work
No	Domestic labor, parenting, household work
	Caregiving
	Craft work
	Volunteer work and civic labor
	Emotional labor, self-disclosure labor

the fields of HCI and CSCW have treated a wide range of everyday life activities, compensated or not, witting or unwitting, as labor.

Driven by the evolvement of computer-mediated labor in HCI and CSCW, we define data labor as encompassing both witting labor activities such labeling images on Mechanical Turk and uncompensated, unwitting ones such as producing content on a social network. Moreover, many widely studied instances of computer-mediated labor are data labor: crowdsourcing [**Sannon and Cosley2019a**], peer production [**Geiger and Halfaker2013a**], and content moderation [**Dosono and Semaan2019a**] are all data labor that advances computing technology and, thereby, benefits computing companies financially. As such, data labor can be seen as a subset of computer-mediated labor.

Digital Labor

Digital labor is another term that has been used to refer to monetized online activities, regardless of whether they occur at traditional workplaces or whether they are compensated

[**Scholz2012**]. In particular, terranova2000free argued that “the Internet is animated by cultural and technical labor through and through, a continuous production of value that is completely immanent to the flows of the network society at large” [**Terranova2000a**]. This work has subsequently inspired in-depth examinations of how online interactions generated value with the commercialization of the internet, e.g. interacting with YouTube video[**Postigo2016**] and managing communities [**Matias2019a**].

While in most cases, prior work on digital labor is relevant to data labor as we've defined here, there may be some examples of digital labor that are not data labor. For instance, cultural production activities like writing fan fiction and private communication activities like private messaging do not necessarily result in improvement of a technology or capital creation. It is worth noting that as developers of large language models such as GPT-3 [**Brown et al.2020b**] and Deepmind's “Gopher” [**Rae et al.2021**] collect massive amounts of content from the internet as training data, more activities will become data labor (e.g. a piece of fan fiction that was previously not data labor may be scraped into a language model training dataset). Furthermore, not all types of data labor are likely to be considered digital labor. Passively produced data such as location data, traffic patterns, and private preference information actively play a role in the improvement of advertising models, navigation algorithms, and commercial recommender systems – however, they are not commonly seen as digital labor [**Scholz2012**].

Crowdwork

Crowdwork is a subcategory of computer-mediated labor in which crowds of distributed laborers complete small-scale tasks for payment. Examples include completing image labeling tasks that enabled the creation of ImageNet [**Deng et al.2009**] and producing texts used to train spam filters [**Ott et al.2011**]. Many instances of crowdwork are unambiguously data labor, although there may be exceptions: completing behavioral experiments on Amazon Mechanical Turk ran by academic institutes–a prominent type of crowdwork[**Hara et al.2018a**]–does not always generate data that leads to an improvement of technology or capital generation, and therefore, may not be data labor.

Data Work in Data Science and Machine Learning

Data work is a relatively new term that emerged to describe data generation, labeling, and cleaning activities for supporting data scientists and machine learning developers [**Sambasivan et al.2021, Møller et al.2020, D'ignazio and Klein2020**]. As machine learning models and other intelligent technologies such as recommender systems become more and more pervasive, scholars have called for attention to the data work that powers these models. In particular, d2020data called for increasing the visibility of data work in the field of data science, so that those performing this work are recognized and valued [**D'ignazio and Klein2020**].

Like digital labor, the term "data work" is very close to data labor, so we will focus on differentiating how the terms have been used. Sambasivan et al. referred to data work as data collection, labeling, analysis, and cleaning, but primarily focused on the activities by those who are "data science workers" (a role discussed at length by Zhang et al. [**Zhang, Muller, and Wang2020**]). Thus, what we refer to as data labor here is primarily the upstream activity of generation and collection, not downstream activities like transforming a column and writing database queries.

Frameworks of Power and Data Leverage

To empower data labor, it is critical to adopt an appropriate definition of power, and what it means to shift more power to data producers. The concept of "power" is a complex, hotly debated one [**Hardy and Leiba-O'Sullivan1998**]. We draw on schneider2018empowerment's work on HCI and empowerment that is intended to "add structure and terminological clarity" to the notion of empowerment in computing research [**Schneider et al.2018**]. Through a synthesis of studies on empowerment, Schneider et al. highlight the distinction between two notions of power: power-to and power-over [**Schneider et al.2018**]. Power-to corresponds to an individual's "ability to do something" (Schneider et al. drew on Arendt's work for this definition [**Arendt1958**]). Applying this notion to data labor, power-to means that people can freely make decisions around their data labor, i.e. choosing not to participate in activities that are data-generating or deleting the data resulting from this labor. Power-over refers to "the relation between multiple actors" [**Schneider et al.2018**] or in Dahl's words, "A has power over B to the extent that he can get B to do something that B would not otherwise do" [**Dahl1957**]. In the context of data labor, power-over means that people can influence those who currently benefit from their labor, i.e. technology operators, around decisions regarding data and data-driven technology.

vincent2021data's framework of *data leverage* further highlights connections between power-to and power-over [**Vincent et al.2021a**]. Data leverage describes how data producers may influence technology operators through three "levers": "data strikes" and "data poisoning" harm a technology operator, while "conscious data contribution" can boost up an alternative operator. Each data lever requires collective action to be effective in which individuals engage via activities like withholding data or manipulating data as a group. Performing these specific actions requires "power-to" (e.g. a legal or legal guarantee that users can delete their data contributions). Power-over is achieved only when a critical mass of participation is reached. In concrete terms, different types of data labor will allow for different abilities (power-to) and may facilitate or inhibit meaningful collective action (power-over). Below, we will describe how each dimension interacts with power-to (the building blocks of data leverage) and power-over (the potential impact of collective data leverage).

The Dimensions of Data Labor

We describe five key dimensions of data labor: *visibility , end-use awareness, collaboration requirement, openness, and replaceability.* While each dimension is a spectrum, we provide examples of data labor that fall on the ends of each spectrum to show how even a relatively dichotomous understanding of each dimension can provide immediate usable insights. We then provide an assessment of how data labor's position along a dimension is related to power and discuss potential opportunities to empower data laborers who are at different positions along the dimension.

Process of Dimension Construction

Our research team collectively identified prominent examples of data labor based on prior work and chose core dimensions that can characterize these examples. We drew from the data leverage framework mentioned above [**Vincent et al.2021a**], arrieta2018should's call for rethinking data generation [**Arrieta-Ibarra et al.2018**], and related discussions [**Jones and Tonetti2020, Bergemann, Bonatti, and Gan2021, Vincent, Hecht, and Sen2019**] to identify what data-generating activities were the focus of the data labor proposal. We then collectively mapped out what dimensions may be useful to characterize these data-generating activities and other prominent ways data producers contribute to data-driven technologies. Our research team then collaboratively merged, ranked, and chose the dimensions that can clearly delineate the wide range of data labor activities happening in real life. As mentioned above, these dimensions are not meant to be exhaustive and given the ever evolving landscape of data-driven technologies, we expect more dimensions to emerge as the public's data labor comes in new forms or is used in novel ways by technology operators.

Visibility: Is the capture of data labor publicly visible?

Data labor can be *invisible* or *visible*, depending on whether the capture of this labor leaves any indication publicly (Fig. 5). When a person's activity leads to some digital records being stored privately with no public digital traces, this is invisible data labor. Examples of invisible data labor include the generation of user interaction logs (for search engines, ad systems, recommender systems) and the production of image labels when users complete reCAPTCHA.

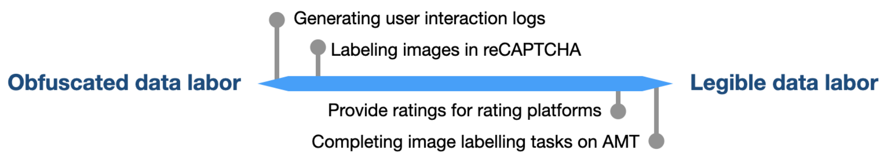


Figure 5. The dimension of visibility: Is the capture of data labor publicly visible? As in all figures in this section, the positioning of data labor instances on the spectrum is approximate and may vary depending on the specific case. For example, for those who understand the background of reCAPTCHA, labeling its images may be visible data labor.
The dimension of visibility.

On the other end of the visibility spectrum exists *visible* data labor – activities that people can clearly see as benefiting technology companies. Examples include generating ratings for rating platforms such as Yelp and Google Maps and completing a crowdwork image labelling task in which the collection of labels is clearly disclosed.

Power: Visibility is positively associated with data laborers' power in their relationship with technology companies. When people do not realize that they are performing data labor, this naturally inhibits their power to withhold or change this labor. This invisibility further limits opportunities for collective action (e.g. a data strike [**Vincent, Hecht, and Sen2019**]) or other types of action (e.g. calling on regulators). In contrast, people performing visible data labor are most likely equipped with more power-to than the invisible condition. They may simply choose not to volunteer or complete a task, so they do not produce any data labor as seen in various non-use cases [**Baumer2018b, Li et al.2019b**]. There is also early evidence about those that perform visible data labor exerting their power-over over technology operators through collective action. For example, Reddit volunteer moderators collectively negotiated with Reddit for better moderation tools [**Matias2016c**] and crowdworkers leveraged Turkopticon to improve their working conditions on Amazon Mechanical Turk [**Irani and Silberman2013a**].

Empowering invisible data labor: Mitigating invisibility – i.e., moving invisible data labor to the visible end of the spectrum – is a first step towards empowering invisible data labor. Given that increasing transparency of labor value helps workers' collective negotiation with employers in traditional labor advocacy [**Khovanskaya et al.2019a**], it is likely effective to apply this tactic to invisible data labor. Researchers and activists may develop tools that measure and communicate the economic or utility value of data labor, such as the Facebook Data Valuation Tool (FDVT), a tool that calculates the worth of Facebook users' attention in real time [**González Cabañas, Cuevas, and Cuevas2017b**]. By making data visible, these tools can better equip data laborers to effectively organize to negotiate with technology companies. Activists can explore how to leverage existing technologies that disrupt the collection of data to make invisible data labor more visible. Currently, privacy-preserving technologies such as anti-tracking browser extensions and protest-assisting technologies such as ad blockers are actively preventing data labor from being used by tech companies and have gained a considerable user base [**Mathur et al.2018b**]. Activists may experiment with providing add-on features to these technologies to highlight the "cost" or "lost ad revenue" users have caused to highlight the value of data labor (e.g. [**Li et al.2018a**]). Additionally, policymakers can play a key role in empowering invisible data labor by requiring technology companies to disclose their capture of data labor. Such efforts could be built upon and extend existing legal frameworks that regulate algorithmic transparency [**gov2019**] and data privacy [**ccp**]. For example, similar to the European Parliament's recommendation standards for algorithmic transparency, policymakers may start designing standards for companies' communication about the process of data collection.

It is worth noting that making invisible data labor visible does not always translate to power for data producers. They may not always have the power-to, i.e. control over their data labor due to external social constraints. For instance, when faced with suspicious, privacy-violating requests, crowdworkers, who are aware they are performing labor, may still complete them because of their need for extra income [**Sannon and Cosley2019b**].

Empowering visible data labor: Those performing visible data labor can immediately benefit from research and tools that strengthen their power-over through collective action. Past research has shown through observation [**Matias2016c**] and through simulation [**Vincent, Hecht, and Sen2019**] that if users collectively withhold their data labor, i.e. data strikes, they can negatively affect data labor-dependent platforms. Given this potential, researchers may further study how to help data producers organize collective action to influence technology operators, answering questions around how different conditions and techniques affect data producers' participation and how to create revenue loss or performance loss for technology operators.

Given the nature of visible data labor, those who perform this type of labor stand to benefit immediately from policies that grant them stronger control over the data output. For example, jurisdictions that have already passed privacy regulations such as GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act) can expand the types of data covered by regulations, i.e. from personal data to other types of data such as monetized user-generated content and crowdsourced datasets of image labels.

End-Use Awareness: Do Data Laborers Understand How Labor Is Used?

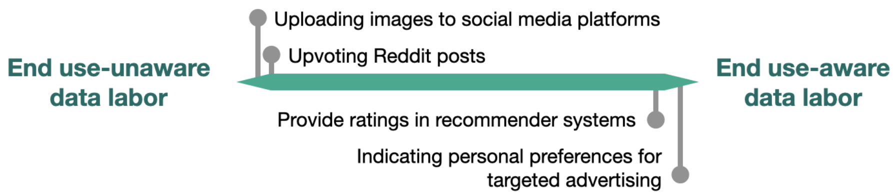


Figure 6. The dimension of end-use awareness: do data laborers understand how labor is used? The positioning of data labor instances on the spectrum is approximate and may vary depending on the specific case.
The dimension of end-use awareness.

While visibility characterizes whether the *capture* of data labor is visible, end-use awareness characterizes the degree to which they are aware of how the resulting data is *used* downstream (Fig. 6). One can have full knowledge of their labor being captured, or put another way, perform fully visible data labor, but have no awareness of their labor's impact. Currently, a variety of data-driven technology companies as well as academic disciplines, including HCI, computational social science, machine learning, and health research, benefit from data labor without the laborers' knowledge. Examples include Reddit content [**Baumgartner et al.2020a, Chancellor, Baumer, and De Choudhury2019**], Flickr images [**Solon2019**] , Wikipedia content [**McMahon, Johnson, and Hecht2017a, Vincent, Johnson, and Hecht2018**], and more. For example, a survey of Twitter users from Fiesler et al. suggests that the majority of users do not realize their content (i.e. tweets) is being used for research [**Fiesler and Proferes2018**].

End-use awareness may change drastically over time as technology operators keep identifying innovative ways to utilize and monetize data labor. In other words, end-use awareness is naturally dependent on what end-uses exist. For example, users who uploaded their images to a social media platform may have high end-use awareness at the time of their uploading action, i.e. knowing their content will help the platform attract more web traffic; but this awareness may decrease if the platform later uses the images for other purposes such as training computer vision models without informing the user. Moreover, end-use awareness may vary from person to person; for example, for the act of uploading images to social media platforms, a computer vision researcher has very likely more end-use awareness than an average user does.

Examples of end use-unaware data labor include producing computer code shared on GitHub that has now been used to train a powerful "AI programming assistant", GitHub Copilot [**Chen et al.2021**], publishing images on Flickr that have been used for facial recognition technologies [**Solon2019**], and upvoting Reddit posts that are later used for large language models [**Radford et al.2019**].

Moving to the other end of the spectrum, end use-aware data labor can be seen in scenarios in which data laborers are directly affected by the technology to which they contributed labor and, therefore, have sufficient knowledge about how the output of their labor is being used. Examples of such technologies include targeted advertising, personalized newsfeed algorithms, and recommender systems.

Power: In general, end-use awareness is likely to lead to more power to data labor. If one knows the downstream impacts of the data they help generate, it is possible to alter such labor purposefully. For example, users reported that they would be demotivated if their contribution to Wikidata, a structured, public access database analogous to Wikipedia, were to be used primarily for profits [**Zhang et al.2022**]. Collectively, a group may have the power-over to shape how these data-driven technologies are developed and deployed. For example, whitney2021hci showed how increasing end-use awareness of data labor assisted community organizations in influencing the City of San Diego's decision on deploying smart city technologies [**Whitney et al.2021**].

Empowering end use-unaware data labor: While it is understandable that technology operators have incentives to keep their data-dependent technologies proprietary and reveal few details (i.e. what data they use, and for what purposes), this practice tends to reduce end-use awareness and, therefore, disempowers the public. Thus, a key direction for researchers and activists is to increase end-use awareness where possible. For instance, such a tool might tell a Wikipedia contributor that their edits on a particular article appeared in Google's knowledge panel. Policymakers can also play a role by mandating end-use awareness, particularly for sensitive data such as biometric data and personal information. Future policies on data use may focus on requiring technology companies to disclose how such data will be used downstream.

Relatedly, understanding how moving end use-unaware data labor to the other end of the spectrum may affect data labor-dependent technologies is also a fruitful area for research. Data producers' concerns about the end use of data labor may disincentivize the production of data labor. As such, increases in end use-awareness may inadvertently reduce the utility of technologies that are currently providing enormous benefits to the public, e.g. Wikipedia

[**McMahon, Johnson, and Hecht2017a, Vincent, Johnson, and Hecht2018**], mental health communities [**Chancellor et al.2019b**], and review platforms [**Li and Hecht2021**], a potential risk that warrants further investigation.

Empowering end use-aware data labor: For end use-aware data labor, given that there already exists significant public use awareness of certain types of data labor in targeted advertising and social media[**Center2019**], researchers and activists may be interested in focusing on these types of data labor and investigating how to transform current end-use awareness to collective action. For example, activists may consider developing tools that make straightforward how end use of data may be affected negatively by data producers collectively withholding or poisoning data. Specifically, such tools may draw from proof-of-concept studies in HCI such as AdNausem [**Howe and Nissenbaum2017b**] and Out of Site [**Li et al.2018a**] and illustrate the downstream effects of data strikes or data poisoning (e.g. "deleting your data would incur loss of ad sales to Facebook"). Of course, end-use awareness need not translate to collective action, if data laborers are content with the current data usage practices.

Additionally, future research should further explore what specific tactics may be of use in facilitating collective action among those who are aware of data labor's end use. By working with community organizations, whitney2021hci highlighted tactics effective for influencing local governments' decision-making about data-driven technologies such as conducting independent data analysis and contesting claims [**Whitney et al.2021**]. Researchers and policymakers may extend their work and examine what tactics may be impactful in the private sector. Researchers and policymakers may also consider strengthening existing legal frameworks that allow users to limit how data is being used downstream.

For example, users may be able to choose clauses such as "Do not use for surveilliance" in addition to "Do not share with third parties" as part of the terms of service for the data they produce [**Contractor et al.2022**].

Collaboration Requirement: Does Data Labor Involve Coordination?

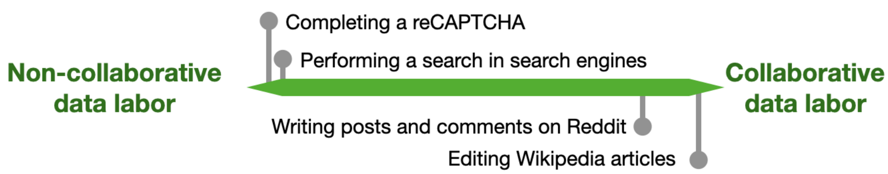


Figure 7. The dimension of collaboration requirement: does data labor involve collaboration? The positioning of data labor instances on the spectrum is approximate and may vary depending on the specific case.
The dimension of collaboration requirement.

Data labor activities can be mapped onto a spectrum from non-collaborative to collaborative, based on the extent to which data laborers work together (Fig. 7). This dimension is informed by the distinction between team work and individual work in computer-mediated labor. The first wave of HCI sought to improve how users complete tasks on desktop computers individually, while the second wave of HCI shifted its focus on "groups working with a collection of applications" [**Bødker2006**]. Similarly, in the context of data labor, data-generating activities can be performed in isolation or involve communication and coordination with other people.

Non-collaborative data labor activities are those that data producers perform in isolation, often with tasks being assigned by technology companies, such as completing a reCAPTCHA and completing data labeling tasks on platforms like Amazon Mechanical

Turk. Other examples include using a search engine or recommender system, as data laborers individually generate data on their own.

When users' data-generating activities involve elements of deliberation, communication, and other forms of teamwork, this is collaborative data labor. This type of data labor is prevalent in social computing systems in which data laborers actively communicate with each other and make decisions about division of labor themselves. As such, technology companies play a less prominent role in shaping data producers' activities than the non-collaborative condition. For example, writing posts and comments, a prominent form of data labor that benefits Reddit (and operators of language models such as GPT-3 [**Brown et al.2020b**]), is largely driven by data producers' intrinsic motivation to produce content and interact with other community members. The posts Reddit users produce depends on their individual interests and background; Reddit as the technology operator has little direct say over such outcome.

Power: The relation between power and collaboration requirement is complicated by the fact that social connections between data laborers can facilitate coordinated collective action but also incur costs to withholding data labor. Data laborers who perform non-collaborative data labor can easily stop performing their tasks without worrying about losing connections with their peers or endangering a collaborative project. However, because of the lack of collaboration in this data labor, users lack shared identity, making it difficult to organize collectively and gain power to influence technology operators. Conversely, those who perform collaborative data labor can theoretically leverage their network for collective action against their "employers", i.e. technology companies. For example, historically, Reddit users coordinated their exit from the platform due to disagreements with the platform's changing policies [**Newell et al.2016**]. However, they may face social cost to exert power-to by withholding or changing their labor if those in their close network are not doing so simultaneously. For example, a Wikipedia editor who does not want their data labor being exploited by for-profit companies and contemplates leaving Wikipedia may fear losing connections with their community members.

Empowering non-collaborative data labor: One step towards empowering non-collaborative data labor is to create connections between users to pave the way for collective action. In traditional labor organizing and crowdworker organizing, workers benefited from having a shared professional identity to pave the way for collective action [**Gray and Suri2019**]. As such, to lay the ground work for labor empowerment, researchers and activists may explore when it is possible to foster a sense of community among users who perform non-collaborative data labor such as Amazon product reviewers. Moreover, activists may benefit from prior HCI research on overcoming challenges associated with collective action by a dispersed, or very loosely connected labor force. For example, by studying crowdworkers' collective labor advocacy efforts, Salehi and colleagues identified two key issues–losing momentum and community frictions–and made corresponding suggestions for design to mitigate these issues [**Salehi et al.2015**]. This prior work can serve as an exemplar to inform future research on how non-collaborative data labor can be effectively organized to advance data producers' shared goals.

Additionally, policymakers may explore regulation proposals that grant data producers collective ownership of the data they produce. Currently, the individual voice and concerns of those that perform non-collaborative data labor have particularly little bearing on companies' practices and decisions. By establishing collective ownership, data may be governed more democratically by those that produce the data labor. One proposal for doing so comes from the California Data Dividends Working Group's recommendations: those that perform non-collaborative data labor may benefit from the establishment of "data relations board", public entities that are tasked to advocate for public interest when companies use massive data labor [**Feygin et al.2021**].

Empowering collaborative data labor: For collaborative data labor, researchers and activists may explore how to grant data producers greater control over their labor while still contributing to their teams and communities. This may be achieved by building alternative technologies that allow users to migrate but stay connected with the old technology in some way. Prior work by fiesler2020moving on fandom communities' platform migration have provided some concrete guidance on this direction [**Fiesler and Dym2020**]. Specifically, they have recommended that alternative technologies allow cross-posting and support data import by users who have had extensive history and interactions on their previous technologies. Policymakers may also be able to play a meaningful role in empowering data labor by mandating or otherwise supporting data portability so data laborers could more easily travel across technologies with their communities.

Openness: Is the Data Resulting from Data Labor Open for Use?

Data labor's *openness* is characterized by how accessible the downstream data is to the public (Fig. 8). This dimension is informed by how computer-mediated labor and, in particular, labor that supports open source projects [**Geiger, Howard, and Irani2021**]

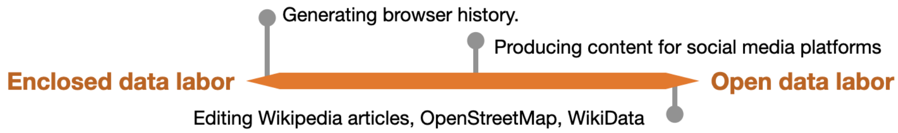


Figure 8. The dimension of openness: is the data resulting from data labor open for use? The positioning of data labor instances on the spectrum is approximate and may vary depending on the specific case.
The dimension of openness.

creates open downstream products. Openness in data labor ranges from being completely enclosed to completely open. Data labor captured by enclosed systems to benefit specific individuals or groups and excludes others is enclosed, whereas data labor in systems that adopt copyleft licenses and make the fruits of data labor public is considered open (e.g. Wikipedia[**McMahon, Johnson, and Hecht2017a**], OpenStreetMap [**Anderson, Sarkar, and Palen2019**], and WikiData). Generally, openness of data labor is determined by the sociotechnical systems in which labor occurs, though researchers and policymakers may make data generated in some enclosed systems open, such as Amazon reviews [4] and ride-hailing records.

Examples of enclosed data labor include generating browser logs, querying information in search engines, and moving around with a device that records location data. In each of these cases, the resulting data output is guarded carefully by technology operators.

Examples of open data labor can be widely seen in academic computing research. For example, the Pushshift Reddit dataset, which provides open access to Reddit content, was used by researchers to create automated moderation tools and language models [**Baumgartner et al.2020a**]. Other prominent examples of datasets created by open data labor include the Movielens dataset

[4] https://nijianmo.github.io/amazon/index.html

[**Harper and Konstan2015**], Wikipedia dumps, ImageNet [**Deng et al.2009**], OpenStreetMap datasets [**Anderson, Sarkar, and Palen2019**], and U.S. census records.

Power: It is possible to gain power-to for enclosed data labor because public access to enclosed datasets is limited. For example, once users have permanently deactivated their Facebook accounts or deleted ratings they have published on review platforms, technology operators will lose at least part of the value from such data labor. Regulations such as GDPR in the European Union [**Por2018**], the CCPA in California [**ccp**], and a variety of similar initiatives have laid the groundwork for the public to exercise their power-to, e.g.removing data from technology operators' records.

Moving to the other end of the spectrum, open data labor, by design, has pro-social goals such as making knowledge accessible to all and incentivizing innovations; however, this openness makes it difficult for data laborers to control who can benefit from their work. Once the aggregated datasets are made publicly available for download, those who produced this open data labor have little power-to exclude technology companies from benefiting from their labor. Even in situations in which users can request to delete their data (e.g. the Pushshift Reddit API), their requests are unlikely to affect all downloaded copies of open datasets that are being re-purposed by private technology companies. At the extreme, when a whole dataset is retracted, technology operators, practitioners, and researchers may still have access to and use its copies, without needing permission or seeking input from those involved with the creation of the dataset (see [**Peng, Mathur, and Narayanan2021**] for an overview).

Empowering enclosed data labor: There have existed extensive efforts from the research community, activists, and policymakers to make more privately-held data publicly available, i.e. transferring enclosed data labor to open data labor. Exemplar outcomes include University of Michigan's data archive[**ICP**], New York University's ad observatory [**NYU**], and City of Chicago's repository of public ride-hailing records [5]. These are meaningful steps towards highlighting enclosed data labor; however, researchers and activists need also to pay attention to how to make data open without sacrificing power-to, that is, how to give individual users control over the data they contributed to such public data repositories.

Data cooperatives [**Pentland and Hardjono2020**] are particularly well suited to empower enclosed data labor given the potentially sensitive nature of the output data. Such proposals help members of the public to gain power-over technology companies via collective bargaining about data usage. Moreover, this approach does not require making the outcome of enclosed data labor completely public, and, therefore, preserves data laborers' power-to if they wish to control how data is being captured and used downstream. Additionally, policymakers may establish other channels through which data producers can collectively and democratically shape the future of their labor. Similar to how shareholder meetings are required for publicly traded companies, policymakers may mandate operators of data-driven technologies to have public channels of communication with users who produce enclosed data labor for them.

Empowering open data labor: To empower open data labor, we first need a more comprehensive understanding of the myriad ways open datasets power private technologies. Gaining this knowledge will help to identify which open datasets underpin today's digital infrastructure and therefore inform activists and policymakers

[5]https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p

of what kinds of open data labor can be potentially leveraged. There existed several studies that have laid the foundation for this research direction, using Wikipedia datasets. Specifically, researchers have studied how Wikipedia data benefits Google Search [**Vincent and Hecht2021, McMahon, Johnson, and Hecht2017a**], Reddit, and StackOverflow [**Vincent, Johnson, and Hecht2018**], and commercial websites in general [**Piccardi et al.2021a**]. As Wikipedia datasets become commonly used in large language models, researchers may further investigate other economic and social benefits of Wikipedia editors' open data labor. Researchers may also be interested in broadly exploring other prominent open datasets whose implications on the tech industry have not been extensively investigated, e.g. OpenStreetMap [**Anderson, Sarkar, and Palen2019, Veselovsky et al.2022**].

The tension between openness and technology companies' value extraction from labor calls for extensive efforts from activists and policymakers to build mechanisms that can support data openness but also give users the rights to the fruit of their labor. Specifically, there exists an opportunity for activists and policymakers to collaborate and build data licensing infrastructure [**Contractor et al.2022**] that let users decide how their data labor could be used in the future at the time when the labor is performed. This would be similar to open software licenses that allow software to be "freely used, modified, and shared" [**osl**] or Responsible AI Licences [**aaa2020**] but give users the flexibility to set constraints on who can use this data in the future.

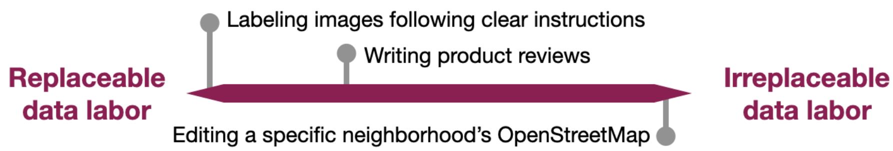


Figure 9. The dimension of replaceability: Is data labor replaceable? The positioning of data labor instances on the spectrum is approximate and may vary depending on the specific case.
The dimension of openness.

Replaceability: Is Data Labor Replaceable?

Like computer-mediated labor, data labor may require certain background, knowledge, skills, or contextual aspects to perform, making it irreplaceable (Fig. 9). Replaceability can also be thought of similarly to the conceptions of skill in labor economics that emphasize modeling workers "endowments" to perform certain task [**Acemoglu and Autor2011**].

There are many examples of data labor that are highly substitutable in the sense that many people, with a variety of backgrounds and contexts, could perform the data labor. Prominent examples include instruction-based labeling tasks, or simple tasks such as fixing typos in Wikipedia articles.

For the majority of data-dependent systems that involve modeling the behavior and preferences of people, model performance will have some degree of specificity to the people who contributed data. A system that models geography will perform better for people in places with more training data [**Johnson et al.2016b**], and a system that models language will likely perform better for languages with more data available.

Power: Data laborers responsible for less replaceable data will have more ability to directly impact technology performance. This means a group of people capable of generating data that is currently underrepresented (e.g. people who can write in languages not currently captured in existing training data) may have more power over those that benefit from their labor. This dynamic is ultimately very similar to how translators who speak a rare language may demand higher wages. From the perspective of advancing responsible AI goals, this dynamic may be useful to leverage. In situations in which technology companies wish to capture the totality of data generated by the public, members of groups currently unrepresented in existing datasets will have more power. It is important to note that the operative decision as to whether data labor from underrepresented groups has above-average or below-average replaceability depends on how technology companies plan to evaluate and deploy their technology (i.e. the choice of "training set").

For data labor that is easily replaceable (primarily, very structured and instruction-based labeling tasks), wielding influence will be harder, and in particular will require larger group sizes. Those who perform easily replaceable data labor (e.g. fixing typos, reporting spam) will only gain influence with exceptionally large-scale collective action.

Empowering replaceable data labor: Researchers and activists may look for ways that help users become irreplaceable, such as identifying and developing unique skills and knowledge as recommended by crowdwork researchers [**Kittur et al.2013**]. Additionally, policymakers should take into consideration the role of societal inequities that prevent data producers from becoming irreplaceable, such as lack of means in gaining digital literacy for certain groups of the public[**DiMaggio et al.2004**]. In other cases, the most efficacious approach may be scaffolding collective action amongst a large pool of people who perform replaceable data labor. For instance, anyone who clicks search engine links is part of the massive pool of search engine trainers. This could be turned into an advantage by building global solidarity around this type of data labor, i.e. in pursuit of a "general data strike" [**Vincent, Hecht, and Sen2019**].

Empowering irreplaceable data labor: Researchers and activists may be interested in leveraging the irreplaceability of certain types of data labor to reduce the power differential between these data laborers and technology companies. This may be achieved by recruiting expert users and niche groups in data strikes against data-driven technologies, such as asking those who are fans of certain movie genres to remove their ratings from recommender systems [**Vincent, Hecht, and Sen2019**].

Discussion

The five dimensions we identified and articulated above are only a starting point for understanding the rich variety of data labor activities. When comparing any two kinds of data labor, there are likely to be subtle differences along each dimension. Critically, these dimensions are immediately useful in that (1) they provide a framework to help people outside of tech companies – such as people performing data labor – to reason about data labor and (2) we can draw on existing knowledge to map out opportunities to organize and empower different kinds of data labor. The dimensions above suggest certain cases in which a data labor lens will be particularly useful to researchers, activists, or organizers; there may be "low hanging fruit", for instance where small design changes can make data labor more legible. Conversely, the dimensions above suggest certain data laborers may be very unlikely to gain power over technology companies, such as those perform open data labor.

While we have painted a broad picture of data labor, many, if not most, computing systems are complex and opaque to the public. As such, there may exist other dimensions that characterize data labor from technology companies' perspective. Below we discuss two such data labor characteristics that are potentially important for technology companies. Additionally, we discuss how researchers may support data labor in their own research practices.

Data Labor from a Technology Company Perspective

The outputs of data labor can take different forms based on how technology companies choose to process and manage it, e.g. databases, digital records, models, predictions. Here, we discuss additional considerations regarding data labor that arise from taking the perspective of a technology company or other data-dependent technology operators (i.e. an organization holding data as capital [**Sadowski2019**]). These factors are important to explore in future work, but compared to the core dimensions above are currently challenging to study because they will require comprehensive knowledge about how data is being collected, processed, managed, and used behind closed doors.

Revenue Generation: The particular causal link between data labor and firm revenue will vary depending on companies' business models. For example, as Google provides API services to directly sell user-generated reviews from Google Maps, the data labor involved with writing reviews is closely tied to the API revenue. In contrast, the data labor involved with using Google Search is indirectly tied to revenue because it has to be processed and aggregated in order to improve search quality, which in turn leads to increased ad revenue. Given the growing use of AI technologies as general enhancements to “free” services (e.g. autocomplete in Google Docs, photo categorization in Apple Photos), there are many cases for which directly tying data labor to revenue will be challenging. Nonetheless, the existence of these technologies suggests the underlying data labor is valuable.

Quantifying data labor’s relationship with revenue will be key for researchers’ and policymakers’ abilities to understand and support the data-generating public. Equipped with methods to assess data’s impact on revenue, researchers and policymakers would be able to empirically assess the economic equity between data producers and technology operators and, subsequently, make informed policy recommendations. Moreover, this framework would also assist technology operators in assessing to what extent their business is being subsidized by data producers and, thereby, gaining a more accurate quantitative understanding of their business. Recent research in this area has made great strides in examining specific revenue streams of data; researchers have assessed Wikipedia’s value to search engines [**McMahon, Johnson, and Hecht2017a**], online communities [**Vincent, Johnson, and Hecht2018**], and commercial websites [**Piccardi et al.2020**]. However, many other instances of data labor remain under-investigated due to the opacity and complexity of the ways data labor generates revenue for companies. One important and urgent area for future work would be to comprehensively quantify data labor’s value for technology companies.

Shelf Life: Organizations that capture data labor may have specific requirements for how frequently new data labor must be captured, i.e. the ‘shelf life’ of data labor. For instance, systems that model real-time variables (e.g. misinformation classification and traffic estimation) need to collect data constantly, whereas some computer vision models may be able to use very old training data (e.g. ImageNet).

An issue with studying 'shelf life' is that public research, by necessity, often uses large datasets without regard to whether the data may have "expired". For example, researchers have provided evidence showing that the Yelp Open Dataset, widely used in location recommendation research, may include data on permanently closed businesses [**Li and Hecht2021**]. In some instances, researchers simply lack access to relevant information about the time frame of datasets. For example, the exact collection time of the BookCorpus dataset, an influential dataset in language models, is hard to determine [**Bandy and Vincent2021a**].

When data labor outputs do not have an expiration date, this is likely to reduce the leverage of data laborers, because it is very difficult in practice to redact a dataset, at least under current laws in most jurisdictions [**Peng, Mathur, and Narayanan2021**]. In recent years, policies that strengthen data producers' right to data, such as GDPR and CCPA [**Por2018, ccp**], have laid the ground work to allow for data producers to delete data if they wish to, providing a venue that is particularly helpful to leverage data with a short shelf life. In the long term, to empower data labor, it may be fruitful to limit for how long data could be used, so data laborers can gain some leverage over technology operators over time. For example, GDPR has mandated that technology operators minimize the time period for which personal data is stored [**Por2018**]. Future policymaking efforts may further expand such mandate from personal data to other types of data. However, the effectiveness of this approach will depend on technology operator choices and details about how often companies retrain models or re-collect data are often kept private. Ideally, with the adoption of practices around the documentation of datasets [**Gebru et al.2018**], data's "shelf life" may also become visible so that data laborers can account for shelf life.

Data Labor outside of Large, For-Profit Technologies

Our road map largely focused on empowering how data labor is valuable for large, for-profit technology companies. However, there exists data labor that generates revenue through improving a technology for other types of entities and organizations such as nonprofits, governmental agencies, and small businesses (e.g. a small store that collects order preferences from its customers to improve its ordering system). For example, open and free knowledge production in Wikipedia allows the Wikimedia Foundation to launch the Wikimedia Enterprise API [**wik2022**]. This API service provides easy access to high volume data with a cost and currently has Google and Internet Archive as key customers. By repackaging Wikipedia editors' data labor, the API generates revenue to support the operation of the system. Future analysis may explore how such data labor outside of our scope may fall onto the five dimensions discussed above to explicate potential ways to empower data laborers valuable for nonprofits and governmental agencies.

Reflections for Computing Researchers

Empowering data laborers means that computing researchers need to be more reflective regarding their own data collection and use. There exists an ongoing discussion on the negative impacts of computing research [**Hecht**]; researchers, across the subfields of computing, e.g. privacy and security, HCI, and AI, have advocated for responsible innovation [**Bates et al.2019, mon**]. Our dimensions pave the way for researchers to pragmatically consider the treatment and recognition of the human labor involved in their innovations. In addition to asking “does my system violate user privacy and autonomy”, researchers may also ask themselves, “does my system recognize and value all labor involved?” One promising direction for researchers to recognize data labor is to include those who generate data in the design and development of data-driven technology, so data producers can play a role in deciding how the fruits of their labor are used. This approach would be similar to algorithm design and governance frameworks in which stakeholders and communities collaborate with researchers, e.g. value-sensitive algorithm design [**Zhu et al.2018**], participatory algorithm design [**Chancellor et al.2019a**], and democratic algorithmic governance [**Lee et al.2019**]. Researchers who currently benefit from data labor may draw from these frameworks and engage with data producers to ensure that their use of data labor is consistent with data producers’ values.

Moreover, as laid out in the dimensions section, there exist ample opportunities for HCI and CSCW researchers to design and test tools to inform and empower data laborers at both the individual and collective levels. This is the key motivation for us to call for HCI and CSCW scholars’ attention to the emerging discussion around data labor.

Conclusion

Through a synthesis of theoretical discussion about data and literature on various types of labor, we introduce a roadmap for empowering data producers by constructing a formal definition and five core dimensions of data labor. For each dimension, drawing on the frameworks of empowerment and data leverage, we prescribe pathways for researchers, activists, and policymakers to empower the millions of people who generate data for large, for-profit technology companies but currently have no say about the governance and use of data.

Chapter III: The Visibility of Data Labor

The invisibility of data labor poses an formidable barrier for data laborers to gain control over their contribution to data-driven technologies. As such, I focus on examining the invisibility of data labor in the next two chapters. Data laborers are crucial to the success of prominent commercial social platforms, such as Reddit, Twitch, and Facebook Groups. Beyond all the publicly visible labor they do generating content, data laborers also perform managerial tasks behind the scenes such as content moderation, fact-checking, and norm-setting. This labor ensures the health and vibrancy of social platforms and is essential for maintaining online communities.

Despite data laborers' utmost importance to many social platforms, they are not always the group that platforms prioritize in design and development, especially for volunteer content moderators. Prominent news outlets have reported that social platforms powered by volunteer moderators such as Reddit and Facebook prioritize revenue-generating user engagement over meeting volunteer moderators' needs [**Peck2019, Post2020**]. Moderators often feel under-appreciated, under-supported, and under-compensated by the platforms that rely on their labor [**Gilbert2020, Matias2016b, Postigo2009**]. This tension between moderators and the platforms they support boils over into public disagreements and disputes, e.g. "blacking out" popular communities on Reddit by making their content private and class-action lawsuits against AOL [**Centivany and Glushko2016, Matias2016b, Postigo2009**].

To properly support moderators and preserve the online communities they maintain, the design and development of social platforms must be rooted in a comprehensive understanding of this labor. Existing approaches to researching content moderation at a large scale focus primarily on moderator activities that leave public visible traces, i.e. removing content and communicating with communities publicly, e.g. [**Chandrasekharan et al.2019, Fan and Zhang2020a, Jhaver, Bruckman, and Gilbert2019a**]. However, new research shows that additional work happens behind the scenes such as managing user behavior and maintaining community settings [**Gilbert2020, Lo2018**]. Without accounting for moderator labor as a whole, developers and researchers of social platforms risk undervaluing and driving away these volunteers and potentially undermining their platforms.

In this chapter, we seek to more completely quantify and characterize moderator behaviors on Reddit. Working with Reddit moderators directly, we collected private moderator logs, called mod logs, from over 900 moderators of 126 subreddits. Private mod logs capture many more moderator actions in addition to the publicly visible ones mentioned above. As such, our dataset allows us to study the work that has fallen through the cracks of prior work and to build a taxonomy of visible and invisible work in content moderation. This broader lens shows that content moderation work is heterogeneous across both the subreddits and across moderators who work on the same subreddit. Moreover, despite being the main area of inquiry for moderation research, comment removal does not paint the full picture of content moderation; it may account for as little as 2% of total human labor across subreddits.

For research efforts on content moderation and online communities specifically, our study complicates prior assumptions about moderator behavior and highlights the limitations of analyzing only visible moderation work. Our work also highlights the richness and scale of the volunteer labor that has helped to enable research and development beyond Reddit by improving user-generated data, e.g. large language models and mental health research [**Brown et al.2020a, Chancellor, Hu, and De Choudhury2018**]. We discuss potential ways for researchers, developers, and labor advocates to understand and support this hidden labor in computing more comprehensively.

Related Work

Human Labor in Computing Systems

Science and Technology Studies (STS) literature has long argued that understanding and supporting workers is the precursor for successful and sustainable computing systems [**Grudin1988, Star and Strauss1999**]. These labor practices have been a core area of interest in social computing [**Geiger and Halfaker2013b**].

For our interests, background labor has been identified as both vital to platform health and simultaneously challenging to study. Background labor is work that is essential for the daily operation and maintenance of systems but is often obscured or ignored by the same systems [**Star and Strauss1999**]; content moderation work is a prominent type of background labor. To understand background labor, researchers commonly use qualitative methods, such as ethnography, interviews, and self-reported survey data. Suchman provided an example of how ethnography around "document coding"—document work completed in a law firm to support attorneys and often misperceived as unskilled, "mindless" labor—unveiled the skills and expertise required [**Suchman1995a**]. More recently, Kriplean called for researchers interested in Wikipedia editor activities to study background labor on the site such as administrative actions and providing social support [**Kriplean, Beschastnikh, and McDonald2008**].

Using content moderation as a case study, we build on the valuable insights in prior qualitative work to characterize background work. Although ethnography provides rich details about work activities, this method does not easily scale to massive remote collaboration across thousands of people. Similarly, interviews and surveys cannot provide granular insights into action-level work activities and also have limitations with self-reporting biases [**Ernala et al.2020**]. , we collaborated with Reddit moderators to collect private mod logs to provide a more expansive picture of their work patterns and practices.

Content Moderation in Social Media

Three branches of the growing literature on content moderation informed our work and guided our analysis.

Invisibility is a known characteristic of content moderation and complicates research of social platforms. Moderator actions are made visible to users through changes to the content of a site. For example, removing a comment will leave traces to non-moderators, because the comment's text will be replaced with "[removed]". Conversely, some moderator actions are not publicly visible on the site. There are limited data traces that signal the occurrence of these actions and corresponding work. For example, when moderators ban users from a subreddit, this action is only visible to the affected users and invisible to the broader community. Qualitative studies have highlighted that much moderation work is made invisible to non-moderators by platform affordances and design decisions [**Gilbert2020, Lo2018**]. This leads to an important observation – what specific actions are not visible to public-facing people (like researchers) and how do they compare to visible work in volume? Given mod logs' expansive coverage of visible and invisible moderator activities, our study lays out a classification of granular visible and invisible actions in Reddit content moderation and quantifies their volume.

Makeup of human moderation work at a granular level is another area of inquiry that has been challenging to study due to the lack of quantitative data about moderators' specific activities. Because comment removal is publicly visible, most community members perceive human moderation work as primarily being content removal [**Myers West2018**]. In contrast, qualitative studies have described the richness and heterogeneity of human moderator labor [**Jhaver et al.2019b, Seering, Kaufman, and Chancellor2020, Seering et al.2019**]. Our work further validates these assumptions and qualitative findings with an action-level analysis of moderator behaviors. Workload is another key area of inquiry in moderation research [**Chandrasekharan et al.2019, Lin et al.2017a**]. The amount of work that moderators perform is hard to quantify due to the invisibility of their activities. To understand what types of moderators face heavier workloads and could benefit from tooling support, researchers have relied on self-reported information and proxy measures [**Matias2019b**]. In our study, we mapped our dataset of mod logs onto each subreddit's posts and comments and in doing so, we provided a quantification of moderators' workload per post and comment.

Background and Methods

Background

On Reddit, each community (called a subreddit) is run by a team of volunteer moderators. Reddit moderators can take many kinds of actions on a subreddit, including approving and removing comments and posts, modifying the visual style of a community, and banning users, among others. All actions that a moderator takes using the built-in moderator functions on Reddit are recorded through moderator logs, or "mod logs". Reddit mod logs are a private record of moderation actions. These logs are only accessible to a subreddit's moderators through the Reddit user interface or the Reddit API. Mod logs are not editable and are updated in real-time as actions occur.

Data Collection

We collected mod logs from two sets of subreddits: 1) a subset of subreddits affiliated with u/publicmodlogs and 2) subreddits recruited by our research team. u/publicmodlogs is a Reddit bot that publishes all mod logs of subreddits to which it is installed. We included 84 subreddits from this list that were active at the time of our data collection, i.e. having at least one user post per day and one user comment per day when we gathered our data between June 2020 and January 2021). Because these 84 subreddits often cover niche topics (cryptocurrency, Not Safe for Work [NSFW] communities, and those with strong anti-censorship views), they may provide limited information about Reddit moderator behaviors more generally. In particular, this dataset does not include any large subreddits. To address this limitation, we directly recruited subreddits to contribute mod logs. We randomly sampled 400 subreddits using Reddit's r/random function and contacted their moderator teams through moderator mail, a private message channel that reaches all moderators on a subreddit. 42 subreddits' moderator teams shared their mod logs to support our research. This set of subreddits included three large communities that have over one million subscribers. We worked with moderators to determine what types of information should be anonymized or omitted during our data collection. We make available our data collection script for those interested in advancing the study of mod logs. This part was reviewed by our Institutional Review Board for human subject research.

Given the sensitive nature of mod logs and the subsequent challenges in collecting this data, it was not realistic to capture a perfectly representative sample of moderators. Instead, we sought to develop a dataset that could catalyze progress towards unveiling and characterizing moderation work that may have been overlooked by researchers and developers previously. We note that in another study conducted using the same dataset, we compared the active moderators in our sample with the whole active moderator population using several publicly available activity metrics such as number of distinguished comments and account age. Although K-S tests show the distributions of these metrics differ between our sample and the population, means and medians are close. And the minimum and maximum values in our sample suggest that it also provides reasonable coverage in values (see Li et al., 2022 for more details). Put simply, our sample is an imperfect but somewhat representative sample of the moderator population.

Dataset Overview

Our final dataset of mod logs includes over three million actions from 126 subreddits and over 900 moderator accounts (including both human and bot accounts). The dataset captured 64 types of moderation actions that go beyond approving and removing content actions and included editing subreddit Wikis or rules, adding flairs to posts, banning users, etc. To avoid confusion, we use the term "moderation" or the verb "moderated" to indicate that one of these 64 action types has been taken on posts or comments. We use the term "removed" to refer to posts or comments being removed by moderators.

	Subscriber count in thousands	***Daily average post count***	***Daily average comment count***	***Data collection span in days***
Mean	350+	70+	700+	142
Max	15,000+	2000+	20,000+	624
75%	200+	40+	500+	169
Median	50+	15+	100+	167
25%	20+	5+	20+	88
Min	5+	1	1	12

Table 6. An overview of our 126 subreddits' subscriber count, activity metrics, and data collection span.)

Dataset overview.

Table 6 provides descriptive statistics about subreddits' subscriber count, daily post and comment counts, and timeframe. To protect moderators' and subreddits' anonymity, we reported all subreddits with an anonymized identifier. This is a combination of a subreddit's rank in subscriber count in our dataset and the category of its topical interest (out of news, gaming, politics, NSFW, and others). For example, r/1humor is the largest subreddit in our dataset and focuses on humor-related content. The mean number of actions per day across the 126 subreddits is 25,812.

Accounting for Invisible Work - A Taxonomy

We began our analysis by determining what work is visible for non-moderators. The outcome of content moderation may be discovered through API services like the Reddit API and Reddit's interface directly such as comment content being replaced with "[removed]". However, many moderation actions are not easily detected by users (if at all), and for those that can be detected, the impact of those changes may fade in a user's memory. For example, changing a subreddit's visual styles would be noticed only by users who recall that there was a previous version of the design; newcomers to a community would not "notice" this change at all. Such changes are not publicly logged anywhere except for the visual style itself, and it is likely that they would be "forgotten".

To help distinguish these levels of visibility, we create a taxonomy of visible and invisible labor. We draw on prior work by social computing scholars in social translucence [**Erickson and Kellogg2000**] and visibility as it applies to organizational systems. Specifically, Treem and Leonardi defined visibility as "the amount of effort people must expend to locate information" [**Treem and Leonardi2013**]. Their definition of visibility is relevant to content moderation because, similarly, technology design affects how visible moderators' work is to others who are not immediately privy to moderation actions. In such systems, accessible information that requires substantial efforts to retain is functionally invisible because people will be unlikely to look for it.

Following this approach, two authors of this paper consulted a content moderation researcher to classify each of the 64 moderation actions in terms of its visibility to non-moderators. Both authors involved are Reddit users and are very familiar with the overall UI of Reddit and APIs. One of the authors is a moderator of a medium-large community on Reddit. The content moderation researcher we consulted is also a moderator on a medium-large community on Reddit and has also worked closely with Reddit moderators.

Actions (their visual representation on Reddit's user interface when applicable)	**Daily occurrences by humans**	**Daily occurrences by bots**
3 - Invisible massive investigative efforts, or simply impossible to know		
Approve content:		
approve comment	1,017	43
approve post	1,159	175
ignore report	334	6
Manage users:		
Ban/unban user (A private message titled "You have been banned from …")	220	20
Mute/unmute user (A private message titled "You have been temporarily muted from …")	24	23
add contributor	5	8
2 – Potentially visible Some investigative effort, e.g. accessing the Reddit API or the Pushshift Reddit API to retrieve all removed posts and querying subreddit wiki pages periodically.		
Remove posts:		
Remove post*, spam post* (Sorry, this post has been removed by the moderators of Moderators remove posts from feeds for a variety of reaso)	821	2,714
Edit flairs/labels:		
Edit flair (This is a flair Post Title)	796	271
Mark nsfw, original, and spoiler (nsfw)	41	48
Change settings:		
Wiki revise	95	481
Set how comments are sorted by default in a thread	4	0
Edit rule	3	0
1 – Easily visible No investigative efforts needed because direct cues are provided through UI		
Remove comments:		
Remove comment, spam comment ([removed])	1,477	4508
Engagement with communities:		
Distinguish	450	3,615
Sticky	235	5256
Lock	137	498

*A post removed or labeled as spam by moderator accounts will still be available for direct visits via URL but its content will be filled with "[removed]" (whereas a post removed by the author themselves will appear as "[deleted]"). The post will also disappear from the subreddit's front page.

Table 7. A taxonomy of invisible and visible moderation actions

A taxonomy of invisible and visible moderation actions by human moderators and bots (some rare actions are omitted for space reason).

Table 7 shows our taxonomy of invisible and visible moderation work on Reddit. Following Treem and Leonardi (2013)'s definition of visibility, the two authors and the researcher collectively mapped all moderation actions onto a three-point Likert scale. The scale corresponds to the amount of effort required for regular Reddit users as well as designers and researchers to know that a moderator action happened. An action is considered invisible (or rated a 3) if it is almost impossible for non-moderators to find any trace or the amount of effort required to get this information is impractical. For example, a user may be able to determine if a post was approved if it had appeared as "[removed]" before and they also remembered it. However, many posts are removed by u/AutoModerator immediately after their submission on the Reddit UI and API, making the task of tracking approvals functionally impossible. As such, we considered the "approve post" action to be invisible. In contrast, an action is rated as visible, or 1, if there are direct affordances in the Reddit UI and API that make the action obvious to all, like distinguishing comments or locking threads. Between invisible and visible actions, there exists a category of actions that are not immediately visible to users and researchers but may become visible with some investigative efforts, which we rate as 2. For example, post removals, although not shown on the front page of a subreddit, can be detected if users and researchers specifically search for removed posts via Reddit APIs or visit the post's URL.

Furthermore, under each level of visibility, the research team clustered actions based on what function they achieve, as also shown in Table 2. Under invisible labor, there are two thematic clusters, 1) approve content—actions that keep comments and posts up, and 2) manage users—actions that determine who could engage with a subreddit. Under potentially visible labor, there are 1) removing posts, 2) edit flairs/labels—actions that assign posts categories but are not clearly labeled as moderator actions to users, and 3) change settings. Under visible labor, there are 1) remove comments and 2) engagement with communities. Because automation is a key strategy for moderators to batch-moderate content, we separate bot actions from human moderators by drawing on prior approaches to bot detection [**Jhaver, Bruckman, and Gilbert2019a**]. We identified prominent bot accounts in our moderator lists such as u/AutoModerator and accounts whose names included words such as "bot" and "auto". In addition to this dictionary-based approach, we also identified extremely active accounts that performed more than 3000 moderation actions in one day, and manually inspected their profile pages to determine whether they were a bot. Many accounts identify themselves as bots in their posting history or profile page. For accounts about which we were uncertain, we contacted subreddits' moderator teams to ask if the account was a bot. In total, we classified 39 accounts as bots out of a total of 967 moderator accounts. Bots accounted for the majority (73%) of the 25,812 daily actions captured in our dataset.

Visible and Invisible Labor

Next, we move to examine the volume of invisible labor; in doing so we test prominent assumptions about moderator labor with this dataset. In this and the following sections, we use the format of stating prominent unknowns or assumptions from prior work and using our dataset to provide new insights or analyze whether the assumption holds.

Who Does Invisible and Visible Work?

Unknown: Qualitative evidence from interview studies has suggested that much of human moderators' work is invisible [**Dosono and Semaan2019b, Gilbert2020**]. However, these findings are based on interviews with moderators from one or a few subreddits and it is unclear whether these findings apply to a large, diverse set of subreddits. Prior work has also found visible traces of bots such as removed comments and removal explanations in comment threads [**Jhaver, Bruckman, and Gilbert2019a, Jhaver et al.2019b**]. However, it remains unclear if bots are used for less visible types of work.

Across subreddits, the share of invisible work for human moderators ranges from 9% to 94% with a median of 43%. Put another way, for half of the subreddits in our dataset, invisible work accounts for no less than 43% of human moderator labor. This quantitative evidence, therefore, supports prior qualitative findings on the invisibility of human labor in content moderation on a much larger scale [**Gilbert2020, Lo2018**]. Conversely, the share of visible work varies from 2% to 68% with a median of 23%. For half of the subreddits in our dataset, visible work accounts for less than a quarter of all human labor. These results suggest human moderators indeed perform a significant amount of invisible work in addition to visible work.

With respect to the visibility of bots' work, we find that across subreddits, bots indeed perform visible work predominantly and are rarely used for invisible work. On average, 56% of bots' actions are visible (Median=62%) and only 6% (Median=2%) are invisible. Bots' focus on visible moderation actions further highlights the need for comprehensive analysis of human labor – if researchers and developers only examine visible work occurring in online communities such as removal, they may inadvertently count bots' work as human labor and overlook most human actions.

Work Makeup

To gain a more granular understanding of moderator labor, we further investigate the makeup of work by bots and human moderators, respectively. Assumption: Bots are primarily used to remove comments and posts and engage with comment threads. Prior work has found that bots predominantly perform two categories of tasks: 1) removing comments and posts and 2) engaging with comment threads through distinguishing and/or "stickying" selected comments (distinguished comments will appear along with a moderator badge and stickied comments will appear at the top of the comment thread) [**Jhaver, Bruckman, and Gilbert2019a, Jhaver et al.2019b**]. However, it is unclear if bots perform any additional work.

Result: Our dataset confirms this assumption and suggests that bots are rarely used for other types of moderation work. Across the total 18,843 daily bot actions (73% of 25,812 actions per day), removing comments and posts account for 38% of actions, and engaging with comment threads is 56% of bot work. The vast majority (94%, 118/126) of subreddits use bots to remove comments or posts, and just over half (52%) of subreddits use bots to automate distinguish/sticky actions to engage with comment threads. While there exist individual incidents of bots automatically updating subreddit wiki pages, this is rare in our dataset.

Our dataset provides additional insights on what types of subreddits have bots working on both content removal and engagement with threads. Notably, the subreddits that use bots for both purposes—content removal and engagement with threads—have a higher median subscriber count (Median=150,000+) than the subreddits that only used bots to remove comments or posts (Median=80,000+). A Mann-Whitney U test indicates that the difference was statistically significant ($U=3079.0$, $p<0.05$). Put simply, larger subreddits are more likely to use bots to automate both content removal and distinguish/sticky actions than smaller subreddits. Taken together, our analysis suggests that bots' use in content removal and engagement with comment threads is especially common among large subreddits.

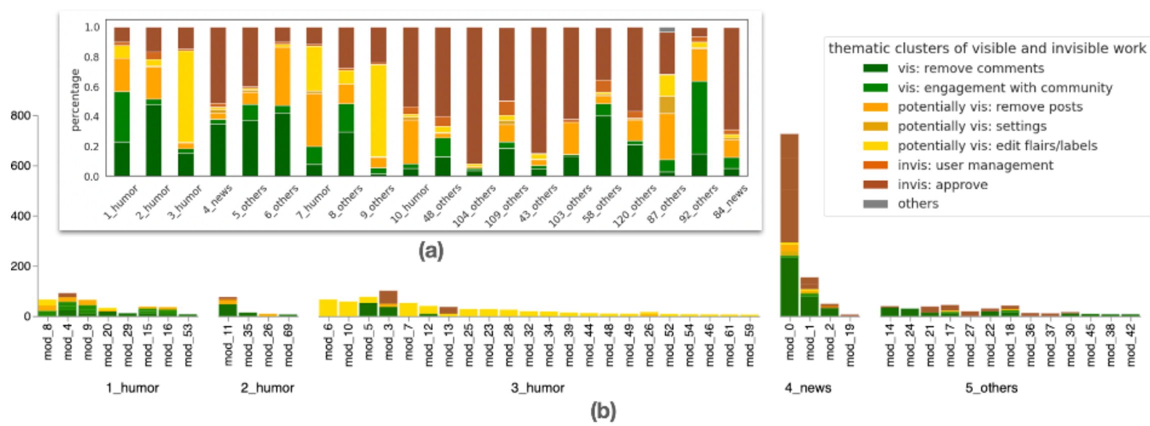


Figure 10. The distribution of human moderation work
The distribution of human moderation work across the seven categories of moderation actions. (a)the distribution per subreddit. (b) the distribution per moderator

Figure 10(a) plots the percentage of each thematic cluster (defined in the Accounting for Invisible Work – A Taxonomy Section) relative to a subreddit's entire human moderation work for twenty subreddits. For a comprehensive overview of all the subreddits' makeup of human labor. The left ten subreddits in the Figure are the ten largest subreddits in our dataset by subscriber count (all with over half a million subscribers). We also included ten subreddits that have the highest volume of moderator activities relative to subscriber count, on the right. We focus on these ten subreddits because their moderator teams are the most active per subscriber and, therefore, offer distinctive insights into heavily moderated subreddits.

Content Removal's Share of Human Labor

Assumption: Content moderation labor primarily consists of comment and post removal. Prior work has shown that community members and the public perceive moderators' main responsibility as removing content [**Myers West2018**]. Public discourse and media reports about content moderation also largely focus on removal and rarely touch on other aspects of this work [**Facebook2021, Wired2014**]. Moreover, much prior work in supporting content moderation prioritizes removal, e.g. [**Fan and Zhang2020b, Jhaver et al.2019a**]. In doing so, researchers and developers inadvertently reinforce the assumption that content removal is the primary component of moderation work.

Result: We find that comment and post removal accounts for 17% (r/9others) to 74% (r/2humor) of human labor among the ten largest subreddits in Figure 10(a) and 2% to 94% across all subreddits in our dataset. On more than half of all subreddits, comment and post removal accounts for less than 61% of overall human labor. These numbers complicate the assumption that comment and post removal is moderators' major responsibility because of how much it varies on subreddits. Furthermore, prior work that used removal-based traces of moderator labor such as removed comments is likely to underestimate "moderation volume" [**Lin et al.2017a**].

Team-Level Heterogeneity

Assumption: Content moderation work is heterogeneous across subreddits. While large social platforms such as Facebook have focused on developing generalizable tools to facilitate content moderation work, researchers have found that moderators of different communities have different values and approaches to their work [**Chandrasekharan et al.2018, Fiesler et al.2018, Seering et al.2019**]. For example, Fiesler et al. (2018) found that communities often express and enforce diverse rules, which imply different moderation practices behind the scenes. Whether this assumption holds has direct implications on what tasks researchers and developers of moderation tools focus on facilitating [**Chandrasekharan et al.2019**].

Result: Returning to Figure 10(a), human moderators engage with diverse types of actions with different emphasis across subreddits. Specifically, each cluster of moderation actions makes up vastly different proportions of overall moderator labor across subreddits as seen in Figure 10(a). For example, approving content, ranked at the top among nine subreddits' human moderators of the twenty subreddits in Figure 10(a), and 43 subreddits across our dataset (out of 126), accounts for as much as 92% in some subreddits' overall human labor, with a median of 34%. Similarly, engagement with community is ranked as the cluster accounted for the greatest percentage of actions on 16 subreddits in our dataset, with a range of 1-78% of human labor (median=6%). Moreover, unlike bots whose actions fall under removing posts and comments and engagement with communities primarily, human moderators cover all seven clusters of actions, regardless of to which subreddits they belong. These findings provide concrete evidence supporting prior work's finding on the diversity of moderator activities across subreddits, e.g. [**Fiesler et al.2018, Jhaver et al.2019b, Seering et al.2019**].

Individual-Level Heterogeneity

Assumption: Moderators of a given subreddit may perform different activities. Prior interview studies with moderators have provided early evidence that moderators take on different roles [**Jhaver et al.2019b, Seering, Kaufman, and Chancellor2020**]. For example, Seering et al. (2020) find a diversity of approaches in moderators' self-described philosophies. Therefore, it stands to reason that moderators may perform different types of actions in their day-to-day practices.

Result: Because larger subreddits tend to have larger moderator teams, we calculate the daily occurrences of the thematic clusters of actions per moderator for the five largest subreddits in our dataset and plot them in Figure 10(b). We find evidence supporting this assumption among these subreddits. On r/1humor, while all moderators remove comments from this subreddit, some also take on other types of work, showing preferences towards removing posts (e.g. mod8 and mod9), some towards approving content (e.g. mod4), and others towards editing flairs/labels (e.g. mod8). On r/3humor, human moderators consistently focus on editing flairs/labels and approving content; however, mod3 and mod5 also remove comments.

Underlying Workload

Content moderation workload is an important metric that can inform future efforts to reduce human labor [**Chandrasekharan et al.2019**]. However, because moderation work leaves limited traces in public datasets, non-moderators have not yet comprehensively measured the volume of moderation work when studying community dynamics [**Lin et al.2017a**]. Prior work has done so with proxy measures, like content removals in [**Chancellor et al.2016b, Lin et al.2017a**]. In this section, we use mod logs to improve our understanding of the amount of work that bots and human moderators perform behind the scenes. We do so by comparing our log data with all posts and comments returned by the Pushshift Reddit API.

Varied Workload for Bots

Unknown: To what extent do bots' workloads differ? In previous work, some human moderators have reported that they emphasize reducing false negatives rather than false positives when using automation tools (i.e., using automation tools to catch as many posts and comments as possible) [**Jhaver et al.2019b**]. However, we have no current evidence about how this strategy plays out across subreddits.

Result: Overall, bots act on from 0% to 96% of posts and 0% to 45% of comments across our dataset. These wide ranges suggest that bots' role in shaping posts and comments varies greatly across subreddits.

Figure 11(a) plots the number of actions bots perform per post (left) and comment (right) among the same twenty subreddits from Figure 10. We find evidence showing moderators' extraordinary attempt to use bots to reduce false negatives on a few subreddits. On r/43others, AutoModerator takes 0.90 remove comment actions per comment submitted. This means that most comments submitted to this subreddit were removed

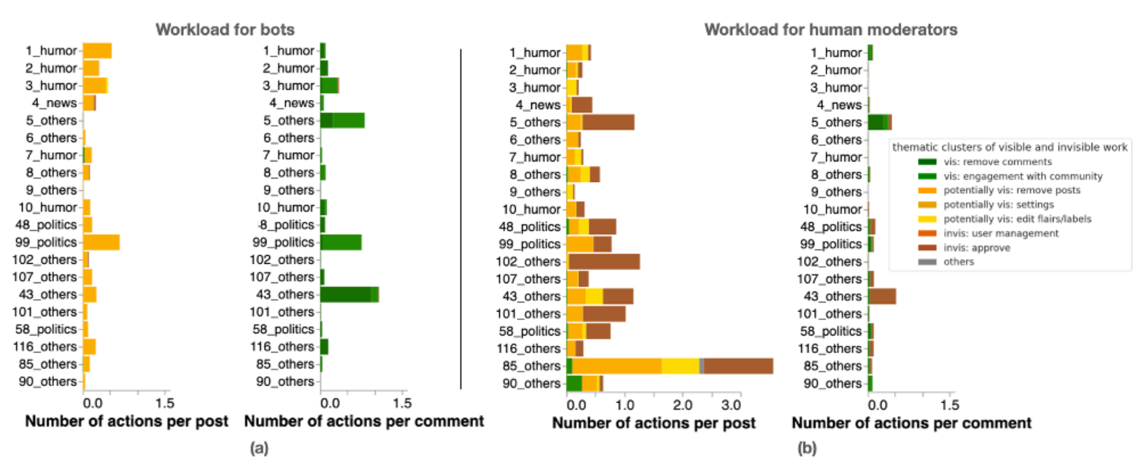


Figure 11. Workload for bots (a) and human moderators (b).

automatically and then manually screened for approval. In extreme cases, moderators may configure bots to manage all posts and comments on their communities.

Furthermore, we find that bots' workload varies between posts and comments; bots performed more actions per post than per comment with a few exceptions (such as r/5others and r/43others). Of the 109 subreddits in our dataset that use bots to moderate both posts and comments, 97 of them have bots performing more actions per post than per comment. We speculate that posts are more prominent than comments on Reddit's UI and, therefore, of higher stakes when moderators configure bots [**Jhaver et al.2019b**].

Varied Workload for Human Moderators

Unknown: To what extent do human moderator teams' workloads differ? Like bots, the workload of human moderator teams' is hard to measure due to a lack of visibility into their actions. This hinders researchers' and practitioners' ability to concretely quantify the amount of human labor involved in supporting online communities.

Result: Figure 11(b) plots the number of actions human moderators take per post and comment, respectively, for the same twenty subreddits. Of the ten largest subreddits, r/5others's human moderation workload is the heaviest, with each comment corresponding to 0.5 human actions and each post corresponding to 1.2 human actions. Across subreddits, human moderators perform, on average, 0.5 actions per post and 0.06 actions per comment. Human moderators have a material influence on posts and less influence on comments. This may be because actions on comments are more limited or that subreddits have more stringent rules for posts than for comments.

Like bots, human moderators' workload varies between posts and comments; humans focus their efforts on posts over comments relatively. On 120 subreddits, human moderators perform more actions per post than per comment. The disparity in workload in Figure 3(b) suggests that with the same amount of increase in posts or comments, different human moderator teams face different amounts of work.

Who Faces Greater Workload?

Assumption: Human moderators of larger subreddits face a greater workload per post and comment. Knowing who performs more work is crucial for researchers and developers to prioritize and meet moderator needs. Prior work has argued that moderator work on larger subreddits is more important because of its potential to affect more Reddit users [**Matias2016b, Matias2019b**]. Research efforts in understanding and supporting moderators also focus on larger subreddits [**Chandrasekharan et al.2019, Jhaver et al.2019b**]. However, we do not know if human moderators of larger subreddits have heavier workloads per post or comment than those from small subreddits.

Result: We did not observe evidence supporting this hypothesis that larger subreddits have more human labor per post or comment. The workload of human moderators per post or comment is not associated with a subreddit's subscriber count (post: spearman's rho = -0.02, $p>0.05$; comment: spearman's rho = -0.08, $p>0.05$). Put another way, large subreddits' moderator teams do not have a heavier workload per post or comment than those of smaller subreddits, though they are likely to have a heavier workload absolutely (post: spearman's rho = 0.64, $p<0.001$; comment: spearman's rho = 0.67, $p<0.001$).

Distribution of Workload

Unknown: How equally is moderation work distributed within a moderator team? Prior interviews have offered insight into the different levels of involvement moderators have with their teams such as "the head mod" vs. "the janitor" [**Seering, Kaufman, and Chancellor2020**]. However, this prior work did not identify how equally they distribute work among themselves. Currently, in data-driven research on online communities, researchers treat all moderators equally by using moderator count in their analysis [**Kiene and Hill2020, Matias2016b**]. However, there may exist moderators who perform little work adding noise to such models.

Result: Returning to Figure 11(b), we find the distribution of moderation work among a subreddit's moderators is highly unequal. We further calculate the Gini index on moderator actions, a measure of inequality, for each of the 36 subreddits with ten or more human moderators. The Gini index values ranged from 0.47 to 0.90 (median = 0.74). Most prominently, in Figure 2(b), r/4_news's most active moderator, mod_0, was responsible for 72% of all the moderation work on the subreddit. Taken together, a subreddit's moderation work likely concentrated on a few moderators, with the rest performing comparatively few actions.

Discussion

Implications on Content Moderation Research

Our findings on the invisibility and heterogeneity of content moderation complicate existing research approaches that rely on publicly available datasets to study moderators' labor activities. As human moderation actions are not always visible, methods that only assess publicly visible work, e.g. removing comments, will very likely leave out a significant portion of work that happens behind the scenes. They may also overlook differences in work makeup and workload across subreddits. Future work must contend with the invisibility and heterogeneity of moderation work if they wish to meaningfully engage with the full scope of moderator labor.

How could research build on our findings? For quantitative studies that characterize moderator engagement, researchers may take our taxonomy as a starting point to investigate invisible labor. Our results suggest that with more investigative efforts into collecting traces of moderation actions, researchers' ability to "see" these actions has the potential to improve accordingly. Additionally, our study suggests that researchers need to ensure the validity of moderator activity metrics across subreddits in their modeling of moderator behaviors given the heterogeneity of moderation labor. For example, metrics that signal strong moderator engagement on one subreddit (e.g. number of distinguished comments) may not work as well on another.

For qualitative work, our findings amplify complementary perspectives about the multiplicity of moderator work that moves beyond content removal [**Gilbert2020, Seering, Kaufman, and Chancellor2020**]. Future work may focus on moderator behaviors that have not yet been fully understood or supported by existing moderation tools. For example, moderators may benefit from tools that automatically approve certain content or users or edit flairs to ease some of their burdens. Our findings suggest that the landscape of content moderation is vastly diverse; no two subreddits are alike. Future work may further explore different ways of content moderation and construct archetypes of moderation strategies.

Implications on Computing Research

Our work highlights the labor supporting the creation of large-scale Reddit datasets and the research communities that rely on Reddit for knowledge production. Reddit data is influential in computing and beyond [**Baumgartner et al.2020b, Bevensee et al.2020**], supporting research on topics such as political extremism [**Chandrasekharan et al.2017, Farrell et al.2019**] and mental health [**Chancellor, Hu, and De Choudhury2018, Choudhury and Kiciman2017**], as well as contributing to powerful machine learning models such as GPT-3 [**Brown et al.2020a**]. When researchers leverage large-scale, user-generated datasets for scientific research, the moderator labor involved in the production and curation of these datasets is poorly understood and documented, even though moderator labor directly influences posts and comments and potentially research outcomes. This documentation (and lack thereof) is especially worrisome when datasets are used for high-stake decision-making [**Bandy and Vincent2021b, Gebru et al.2020, Geiger et al.2020, Proferes et al.2021**]. Gebru et al. have cautioned that without proper documentation of datasets, researchers may make false assumptions of representativeness or generalizability and research outcomes [**Gebru et al.2018**]. Proferes et al. further pointed out that in the case of Reddit, two prominent contextual factors—community culture and demographics—may influence model generalizability [**Proferes et al.2021**]. Indeed, at one extreme, r/43_others' moderators filtered the bulk of the comments submitted to this subreddit and, therefore, greatly influence what content is available on this subreddit. But r/43_others is no singular or novel outlier - as seen in Figure 3, there are several subreddits whose moderators frequently made decisions about what content to remove or approve and thereby, affect their subreddits' content availability. These findings affirm moderators' role in shaping user-generated content and highlight the importance of accounting for content moderation in dataset documentation and research more generally.

Supporting Labor in Social Platforms

Our findings also problematize existing approaches that examine only the visible part of background labor—work that is essential to systems' operation and maintenance but often overlooked by those involved [**Star and Strauss1999**]. Currently, data about the invisible part of background labor is largely held behind closed doors of private companies and is difficult to access for researchers.

Our data collection and analysis point to some potential directions for researchers to resolve these tensions. First, researchers may collaborate with workers directly (like moderators) and deploy tools that collect their log data with strict privacy protection. In the crowdwork domain, tools have been developed to allow crowdworkers to see their hourly wage and simultaneously quantify invisible, unpaid labor for researchers [**Hara et al.2018b, Toxtli, Suri, and Savage2021**]. Future work may explore how this approach could benefit uncompensated digital labor, such as volunteer content moderation and peer production while helping moderators conduct their own "time studies" [**Khovanskaya et al.2019b**].

Second, our results on the sheer volume of volunteer labor necessary to maintain online communities further highlight the importance of recognizing and supporting volunteer labor. Reddit moderators have long needed better support for their work as well as protection against the risks associated with their role such as online harassment [**Gilbert2020, Matias2016b, Matias2019b**]. One factor contributing to their lack of negotiation power in their relationship with platforms is the invisibility of their labor and an inability to quantify their contributions (Li et al., 2022). Designers of computing systems could consider improving the visibility of moderation work to correct these misperceptions and focus internal resources to support invisible work [**Suchman1995b**]. This could be done through interface changes or public reporting such as "this subreddit's moderator team has worked 18 hours for the community in the past week". However, we strongly caution against wholescale attempts at making all invisible moderation work visible given the risks of social surveillance and harassment by bad actors [**Gilbert2020**]. Any attempt that seeks to increase the visibility of moderation work needs to contend with the importance of moderators' privacy, safety, and wellbeing.

Limitation

Although mod logs provide expansive coverage of moderator activities, there still exists invisible moderation work that is not present in mod logs. Two prominent examples are responding to mod mails and deliberation within moderator groups. Prior qualitative work has noted both the importance of this work and the challenges in capturing it [**Dosono and Semaan2019b, Gilbert2020**]. Our study did not characterize such activities given mod logs' limitations. There are other opportunities to understand, characterize, and support these untraceable moderator activities—a fruitful area for future research. One may explore working with moderators even more closely by conducting diary studies to address this limitation.

Conclusion

Using Reddit moderation logs, we complicate prior assumptions about content moderation work and highlight how moderator labor has been partially overlooked or misunderstood. Specifically, we expose the amount of invisible labor in moderation and uncover heterogeneous work makeup and varied underlying workload. Our study highlights the importance of accounting for obscured human labor in content moderation and computing research in general that relies on Reddit data.

Chapter IV: Making Data Labor's Value Transparent

In this chapter, I sought to measure the monetary value of data labor, as a step toward further highlighting the important but often invisible labor force going into data-driven technologies. Data labor underpins some of the greatest technological innovations and advances in recent computing history. In addition to non-profit and open-source initiatives such as Wikipedia and Linux, data producers like online volunteers also supports highly valued technologies such as Stack Overflow (a question-and-answer website sold for 1.8 billion in 2021). Similarly, social media platforms such as Facebook Groups, Reddit, and Discord prominently depend on fleets of volunteer moderators to build and manage communities with millions of users and, thereby, keep these platforms viable [**Gilbert2020, Matias2019b**]. For years, this business model seems fair and plausible because data producers receive free or low-cost services in return, e.g. search engines, online space to communicate with others, and personalized recommendations.

But with data being increasingly valuable and powerful for businesses, questions arise about the legitimacy and fairness of monetizing data producers. Put another way, the business model once seemed to be equally beneficial to data producers may be exploiting data producers. Moreover, volunteer-supported, for-profit technologies set up inequitable power structures in the technology sector. Data producers who provide the crucial labor supporting these companies are subject to worsening working conditions [**Matias2016b**], monetization without consent [**Arrieta-Ibarra et al.2018, Li et al.2019a, Vincent et al.2021b**], and potentially exploitation [**Terranova2000b**]. More broadly, online volunteers often have little power to shape the technology they co-create with for-profit companies [**Vincent et al.2021b**].

From a policymaking perspective, online volunteer work creates new labor mechanisms by subsidizing actual compensated labor [**Postigo2009**], and scholars have suggested that companies profiting from this free work may be contributing to an industry-wide decline in labor share (the proportion of business income allocated to wages) and, subsequently, exacerbating income inequality [**Arrieta-Ibarra et al.2018, Posner and Weyl2018**]

To form a more equitable relationship between technology companies, data producers, and policymakers, we need transparent and rigorous evidence about the value of privatized data production. Without evidence about their work's monetary value, data producers remain uninformed and disadvantaged when seeking to shape the technologies they co-create with companies. This disadvantage has posed a drag on volunteer productivity and business growth historically, as seen in the class-action lawsuit by AOL moderators in 1999 [**Postigo2009**] and the collective protest by Reddit moderators in 2015 [**Matias2016b**]. More broadly, opacity in online volunteer work's value hinders the public's abil-ity to address corporate influence on the technological eco-system that is powered by members of the public [**Vincent et al.2021b**]. Lastly, policymakers, despite making an effort to account for the privatization of online volunteer work in financial regulations [**Au-Yeung2019, Commission2017**], have not yet had sufficient evidence and knowledge to make pragmatic policy recommendations. In short, assessing the value of online volunteer work is a first step toward supporting volunteers, the public, and policymakers in enabling more equitable power structures in the technology sector.

In this study, we empirically assess the value of a particularly prominent type of data production work—Reddit volunteer moderation. Reddit is one of the most visited websites in the U.S., with fifty-two million daily active users

[**Patel2020**], and the website actively plays a role in the public's news consumption [**Stoddard2015**], topical discussions [**Gilbert2020**], and social support for mental health [**Chancellor et al.2019c, Chancellor, Mitra, and De Choudhury2016, Choudhury and Kiciman2017**]. Reddit is organized into thousands of topical communities, called subreddits. Each subreddit is run by its own volunteer moderators, who make daily decisions about community rules, who may participate, and what content will stay online. Although Reddit moderators have reported the benefits of volunteer moderation models such as independence and tailored community experiences, many also experience frustration about their labor not being recognized and supported by the platform they help to maintain [**Gilbert2020, Matias2016b, Matias2016a**].

To assess the value of work contributed by Reddit moderators, we use a novel dataset of private moderator logs ("mod logs") that we collected from 126 communities by working with moderators themselves. This dataset provides more comprehensive coverage of moderation activities than any existing datasets (such as publicly available datasets of removed comments) and allows us to infer the minimum amount of time moderators volunteered. Using linear mixed-effect regression, we estimate that the whole volunteer moderator population on Reddit spent at minimum 466 hours every day performing moderation actions in 2020. Using the median hourly rate among U.S. commercial content moderators on UpWork ($20/hr), we estimate these labor hours amount to 3.4 million USD a year, equivalent to 3% of Reddit's revenue in 2019.

Our work provides the first empirical estimate of the value of data production that powers Reddit, a highly valued social networking site. In doing so, we contribute a better understanding of data producers' role in technology companies' financial success, and, ultimately, help to inform collective negotiation, public debates, and policy recommendations. Additionally, our ability to draw the estimate is unlocked by our novel method that can be generalized to different types of online volunteer work.

Related Work

Online Volunteer Work and Its Impact

Online volunteer work, ranging from open-source projects to peer production to content moderation plays a crucial part in the day-to-day function of prominent computing systems. In recent years, as the output of online volunteer work is repurposed for technological innovation such as language models, online volunteers become a digital labor force that is more important than ever. For example, revolutionary language models such as GPT-3 depend on texts volunteered by Wikipedia editors and Reddit users alike [**Brown et al.2020b**]. Similarly, GitHub's Copilot, a technology that assists programmers in programming is built upon codes published on GitHub. And prominently, many commercial companies and, in particular, cloud computing platforms, benefit from Linux volunteers' work tremendously.

While online volunteer work plays a central role in open-source and commercial computing systems, how to value this work remains unresolved. Researchers have examined how specific outcomes of online volunteer work such as Wikipedia links have benefited technology companies (e.g. [**Heald, Erickson, and Kretschmer2015, Piccardi et al.2021b, Vincent, Johnson, and Hecht2018**]); however, there has not been a way to comprehensively assess the entirety of human labor. Our work takes a first step towards solving this problem by focusing on labor hours, as described in detail below.

Content Moderation

Content moderation work on Reddit is a prominent case of online volunteer work. Reddit relies on its volunteer moderators to manage thousands of online communities to keep its business viable. These moderators perform a wide range of tasks such as setting up community rules, approving content, removing harmful content, and providing explanations for content removal (e.g. [**Gilbert2020, Jhaver, Bruckman, and Gilbert2019a**]). Moderators also regularly communicate with their peers and community members to discuss and shape community norms [**Dosono and Semaan2019b, Gilbert2020**]. However, much of this work does not leave any publicly visible traces on Reddit [**Li, Hecht, and Chancellor2022**]. Moderator logs, a type of private data that is only accessible to a subreddit's moderators, provide an opportunity to more comprehensively capture moderation activities than using publicly visible traces such as removed comments. Although mod logs do not capture all moderator activities, they are a step forward in accounting for the invisible part of moderator labor.

The volunteer-driven approach to content moderation is not the only one employed by social platforms; another approach commonly seen in the technology industry is to hire commercial content moderators who moderate content for compensation [**Roberts2019, Seering2020**]. Compared to the volunteer-driven approach, commercial content moderators have a set of platform guidelines to follow and, therefore, potentially have less independence in managing online communities. However, both groups of moderators perform similar activities; Ruckenstein and Turunen argued that "commercial content moderation has similar aims as community moderation, seeking to support and nurture the online conversation with situated practices" [**Ruckenstein and Turunen2019a**].

Methods

Data Collection

Estimating the amount of time all Reddit moderators spend on moderation is very difficult because there is no publicly available, comprehensive data about moderator actions and activities. Prior work has inferred moderator activity from the amount of public content removed to approximate overall moderation labor [**Chancellor et al.2016a, Cheng, Danescu-Niculescu-Mizil, and Leskovec2015, Lin et al.2017b**]. However, this removal-based approach underestimates the amount of work volunteer moderators complete because content removal is known to be a fraction of their activities [**Gilbert2020, Lo2018**].

To address this major impediment to capturing online volunteer labor, we used private moderator logs ("mod logs") mentioned in the previous chapter. Despite being the most comprehensive digital record about moderator activities, mod logs do not capture the entirety of the work moderators do and provide an estimate of the minimum amount of time spent by moderators. We discuss this limitation below.

Because we are only interested in human moderators in estimating volunteer hours, we removed automated moderator accounts or bots from the dataset, drawing from methods used in prior work [**Jhaver, Bruckman, and Gilbert2019a, Johnson et al.2016a,**

?]. After removing bot accounts, this dataset contains over 800,000 actions from over 900 human moderators.

Estimating Moderation Action Duration

To infer how long moderation work took for each moderator based on mod logs, we followed the process from prior work on Wikipedia session analysis of editor activities [**Geiger and Halfaker2013c**]. Figure 12 provides an overview of our approach. Mod logs only give information on the end timestamp of an action. To estimate how many seconds each action took, we identified "streaks" of actions and calculated how many seconds have passed between each action and its prior action. A "streak" of actions is a series of actions taken sequentially by a subreddit's moderator. Following prior work [**Geiger and Halfaker2013c**], we capped the interval between two adjacent actions' end timestamps at 60 seconds or less. Put another way, if 60 seconds have elapsed between two adjacent actions taken by one moderator, these two actions will be classified as belonging to two separate streaks. Because no prior actions exist for isolated actions or the first action in a streak, we assigned each such action the median value in how long the corresponding moderator spent on performing this type of action in general. Finally, we calculated moderation session duration at the moderator level by aggregating our dataset.

Limitations

Our sample of moderators is not a random sample of the overall moderator population on Reddit. To access one moderator's mod logs from a subreddit, Reddit's API requires

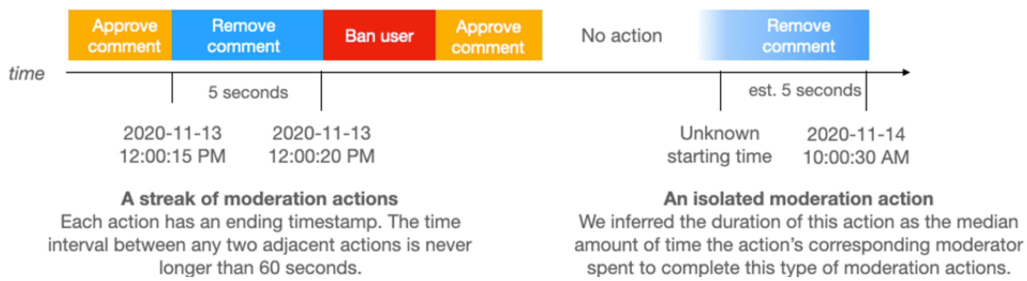


Figure 12. An overview of our approach to assessing how long each moderation action took

our bot to be granted access to mod logs of all the moderators on the subreddit. As such, it is not realistic to randomly sample moderators on the site and collect their mod logs, especially given moderators' privacy concerns about sharing access to this type of data. Below, we compared our sample of active moderators with the whole active moderator population. We found that despite statistically significant differences in several activity metrics such as daily distinguished comment count, our sample's means, medians, and standard deviations do not meaningfully deviate from those of the whole active moderator population.

As mentioned earlier, mod logs do not include all moderator activities. For example, replying to moderator mail and deliberation, two types of moderation work reported in prior work [**Dosono and Semaan2019b, Gilbert2020**] do not appear in mod logs. As a result, our estimate of moderation duration is designed to be a lower bound estimate. Although we provide an underestimate, our work serves as a first step towards quantifying this labor and paves the way for future work to comprehensively quantify moderation hours.

Final Mod Logs Dataset

After collecting mod logs and inferring moderation action duration, we aggregated the dataset to estimate the amount of time moderators spent moderating per day, what we call daily moderation duration. In our sample, daily moderation duration is widely dispersed. This is in part due to a very long tail of inactive moderators – the median in daily moderation duration is 10 seconds (Mean=68 seconds). The number of minutes moderators in our dataset spent every day amounts to 1023 minutes and the top 10% of moderators are responsible for 68% of the daily total, spending between 3 to 40 minutes daily on moderation. The top 20% of moderators spend more than one minute per day and are responsible for 82% of the daily total, which approximately follows the Pareto principle (the 80/20 Rule). For a breakdown of moderation actions, see our in-depth analysis of the makeup of moderator labor using this dataset.

To provide further insights into daily moderation duration, Table 8 provides the correlation matrix for main publicly available activity metrics. These metrics are from the metadata, posts, and comments collected from Reddit's official API service and the Pushshift Reddit API, a volunteer-led repository of Reddit comments and posts used widely in scientific research [**Baumgartner et al.2020b**]. Most prominently, more active moderators, i.e. moderators who work longer daily, are more likely to leave distinguished comments (comments that are posted by a moderator and publicly marked with a badge icon or a "[M]" label) on their subreddits (spearman's rho=0.60, $p<0.001$). They are also somewhat likely to be in subreddits with more posts daily (spearman's rho=0.26, $p<0.001$) and more active moderators (spearman's rho=0.21, $p<0$ .001).

	Daily moderation duration	*Number of subscribers*	*Daily post count*	*Daily post count per subscriber*	*Comment count per post*	*Number of active moderators*
Number of subscribers	0.12***					
Daily post count	0.26***	0.28***				
Daily post count per subscriber	0.06	-0.16***	0.19***			
Comment count per post	-0.01	-0.16***	-0.18***	-0.24***		
Number of active moderators	0.21***	0.41***	0.82***	-0.08*	-0.13***	
Daily distinguished comment count moderator	0.6***	0.27***	-0.02	-0.04	-0.0	0.03

Table 8. Correlation matrix for main publicly available activity metrics. Note: p<0.05; ***p<0.001.

Statistics about Hourly Rates for Comparable Commercial Content Moderation Service

To estimate the value of the hours moderators spent on moderation work, we sought to collect statistics about the hourly "wage" or payment rate for comparable paid work. Currently, official wage data sources such as the U.S. Bureau of Labor Statistics do not offer statistics about commercial content moderators, possibly due to this position being a relatively new occupation. To generate an alternative source of comparable wage data, we turned to UpWork, a crowdwork marketplace in which content moderation experts publish their hourly rates to potential clients. We used UpWork because it is a prominent crowdsourcing marketplace and a frequent subject of research, e.g. [**Foong and Gerber2021, Foong et al.2018**].

We found 160 commercial content moderators on UpWork by searching for keywords related to content moderation and community management services such as "content moderation" and "community manager". We used a new account to minimize personalization in search results. Commercial content moderators' rates varied widely, from 3/$hrto$160/hr, and have a long-tailed distribution (Median=10.0/hr, $Mean$ =14.6/hr, all in USD). Out of the 160 workers, 31 are located in the United States and their median hourly rate is \$20/hr (Mean=\$26/hr). The Philippines, where many U.S.-based social media platforms such as YouTube and Facebook employ many commercial content moderators, has the largest number of workers in our dataset, 50, and the median hourly rate is \$6/hr (Mean=\$10/hr) for this population. Reddit moderation work may differ from commercial content moderation services at a granular level; however, they are somewhat comparable given the two roles sharing similar goals and activities [**Ruckenstein and Turunen2019b**]. As such, the hourly rate for commercial content moderation can serve as a reasonable proxy for volunteer moderation's market value.

Modeling Moderation Duration

To estimate how many hours of moderation work occurred on Reddit site-wide, we built a linear mixed-effect regression model to extrapolate from our sample of moderators to Reddit's overall moderator population. Specifically, we regressed sampled moderators' daily moderation duration onto their publicly available activity metrics and then applied the model to the overall moderator population.

As mentioned above, our sample is not a truly random sample of moderators. Before proceeding to modeling, we first verified that the moderators in our dataset are at least a somewhat representative sample of the whole active moderator population on Reddit such that findings from our sample can be reasonably generalized. We defined active moderators as moderators who left at least one distinguished comment between November 2020 and January 2021 on the subreddit they moderate. This definition is motivated by the fact that distinguished comment provides high recall for active moderators in our sample: the 378 moderators who have left at least one distinguished comment contributed 97% of the actions in our dataset. We identified 21,522 active moderators on Reddit in 2020 using this technique.

	Our sample N=378					**The whole population of active moderators N=21,522**				
	Mean	*Median*	*Std*	*Max*	*Min*	*Mean*	*Median*	*Std*	*Max*	*Min*
Distinguished comment count	1.08	0.23	2.47	17.94	0.02	0.81	0.1	3.49	119.80	0.02
Distinguished post count	0.01	0	0.02	0.32	0	0.01	0	0.06	3.98	0
General comment count	6.11	2.90	9.54	87.67	0.02	6.09	2.38	11.74	250.82	0.02
General post count	0.50	0.07	2.34	40.31	0	0.79	0.08	4.29	229.95	0
Account age in days	2248.14	2311.63	1114.33	5403.98	180.19	2072.25	2050.67	1208.87	5729.53	57.86
Comment karma	92454.24	38433.5	212544.26	3080329	11	61049.88	17499.5	155825.65	4509949	1
Link Karma	104273.18	12638	382020.75	5587774	1	112384.42	9681.5	714638.92	35589509	1

Table 9. A Comparison of Our Sample and the Whole Active Moderator Population

Table 9 shows the comparison between our sample and the population of active moderators using publicly available key activity metrics. The differences in these metrics' means and medians are either insignificant or small—i.e. although the differences are significant, the effect sizes are minimal. However, notably, all two-sample K-S tests reject the null hypothesis that our sample's distributions of these activity metrics follow those of the whole population. Therefore, we expect models derived from our sample would reasonably but imperfectly generalize to the whole population. Put another way, while our sample provides a novel and deep look into moderator activities, this insight requires a moderate sacrifice in sample-population alignment, which is a common limitation in situations in which representative sampling is difficult to implement, e.g. (Killingsworth, 2021).

We regressed daily moderation duration onto publicly available subreddit and moderator activity metrics, using a linear mixed-effect regression model with the 378 active moderators in our sample. We use log-transformed moderation duration as the dependent variable so that the model has normally distributed residuals. The distribution of moderation duration among a subreddit's moderators varies across subreddits. In particular, the Gini index for the 32 subreddits with no fewer than ten human moderators ranges from 0.23 to 0.94 (median = 0.76), suggesting different levels of inequality in workload distribution across subreddits. As such, we chose the best fitting covariance structure to address heterogeneity in residuals (Zuur et al., 2009).

We added a subreddit-specific random intercept to account for the hierarchical structure of our dataset, i.e. moderators being grouped into subreddits. We also added a random slope for distinguished comment count due to the potential variability in its association with daily moderation duration across subreddits. While distinguished comment count is strongly correlated with daily moderation duration overall, the relationship between distinguished comment count and moderation duration may be subreddit-specific. These varied slopes justify a random slope for distinguished comment count per subreddit in our model, which indeed significantly improved our model fit as measured by AIC.

Table 10 shows the coefficients of the explanatory variables from the model, with the dependent variable being log-transformed daily moderation duration. Controlling for other moderator and subreddit activity metrics and for random variability at the subreddit level, with 1% of increase in distinguished comment count, moderators spend 0.8% more time on their moderation duties. This association points to a promising proxy for future research that seeks to understand and model moderation workload on Reddit. Using a linear regression model with distinguished comment count as the only explanatory variable, we observe that this metric explains 44% of the variance in daily moderation duration

	Estimate
Log(Daily distinguished comment count)	0.81 ***
Log(overall comment count / distinguished comment count)	0.05
log(Comment karma)	1.24 *
Log(Link karma)	0.70
Account age in days	-1.28 **
Log(subreddit daily post count)	10.48 ***
Log(Subreddit daily comment count per post)	4.45 ***
Subscriber count	-9.22 ***
NSFW	0.24 ***
Log(Number of active mods)	-0.02 ***

Table 10. Numeric results of the linear mixed effect regression. *p<0.05, **p<0.01, *** P<0.001

across moderators. Future work that seeks to study and identify active moderators may use distinguished comment count as a key indicator.

Results

Applying our model to all the active moderators on Reddit in 2020, we estimate that moderators spent a total of 466 hours per day performing moderation actions.

As a robustness check for our extrapolation, we calculated the sample's weighted mean in daily moderation duration as a proxy to the population mean. We followed the propensity score weighting approach. We first calculated each moderator's propensity score, i.e. the probability of being included in our sample with a logistic regression model. We then used the inverse of propensity score as each moderator's weight. This weighting process yielded a mean value of 80 seconds per day. This weighted mean corresponds to a sum of 359 to 611 hours by the whole population at the 95% confidence interval, encompassing the point estimate derived from our regression model.

Applying Hourly Rates

The value of the 466 labor hours is contingent on the market rate for content moderation. We take two approaches for our estimation. In our first approach, we assume that companies in the market for content moderation services always hire workers who charge the least. In our case, the 59 such workers (466/8, assuming that each worker can work eight hours a day) charge \$3/hr to \$12/hr with a mean of \$8/hr. As a result, this approach leads to an estimate of 1.4 million USD for the 466365 estimated moderation hours, equivalent to 1% of Reddit's revenue in 2019 (120 million USD).

In our second approach, we assume that companies looking to hire content moderators offer a fixed rate. We consider several rates from the UpWork dataset in addition to the \$15 hourly wage strongly advocated by scholars and crowdworers (Rolf, 2015; Whiting et al., 2019) and the U.S. federal minimum wage (given the U.S.'s status as the primary market for Reddit (Statista, 2021)). The value of Reddit volunteer moderators' labor is 3.4 million USD (3% of Reddit's revenue in 2019) if calculated with the median rate of U.S.-based UpWork workers and 4.4 million USD (4% of Reddit's revenue in 2019) if calculated with the mean. Additionally, we consider the Philippines, the country with a prominent labor force of content moderators (Roberts, 2019); the amount of work volunteer moderators completed on Reddit is worth 1 million USD using the median rate of Philippines-based workers and 1.9 million USD using the mean.

Discussion

Our work quantifies the value of labor subsidy volunteer moderators contribute to Reddit, a highly valued technology company. There exist many other prominent, highly valued technology companies and products that rely on volunteer labor, such as Google Maps, Yelp, and Facebook Groups. In this section, we first explore how our results can assist important stakeholders—online volunteers, the public, and policymakers—in ensuring online volunteers are adequately supported and recognized by these companies. We then discuss how future research may further improve this estimate.

Implications for Online Volunteers

Moderators have already voiced their displeasure over poor support for their labor. In 2015, Reddit moderators made thousands of subreddits inaccessible to the public in a protest against inadequate tooling and administrative support - this protest blocked much public web traffic to the site [**Matias2016b**]. Historically, tensions between volunteers and companies have spilled into legal disputes. In the late 1990s and early 2000s, America Online (AOL) moderators filed a class-action lawsuit to dispute the company's management of moderators [**Postigo2009**].

Our estimates highlight potential opportunities for online volunteers and companies to form a better, more sustainable relationship that can maintain the health and vibrancy of the technologies they co-created like Reddit. Our study quantified the important value Reddit moderators bring to the company, and volunteer moderators could highlight this value in their conversation with Reddit to advocate for resources needed to successfully manage online communities. For example, as research and prior historical examples have shown that Reddit's existing moderation tools fail to support volunteer moderators' work [**Matias2019b, Postigo2009, Seering et al.2019**], moderators could use the value we describe here as a talking point to demand software engineering efforts that is equivalent to their collective volunteer hours to improve these tools.

More broadly, knowing the amount of this labor subsidy they supply to Reddit can help volunteer moderators to advocate more strongly for decision-making power in the platform's day-to-day operation such as updating site-wide content policies and division of responsibilities between themselves and the site [**Matias2019b**]. Our finding on the long-tailed distribution of moderation work suggests that a small share of moderators might have particularly strong negotiation power. This raises the question for future work about how (or whether) to ensure that any collective negotiation with Reddit is representative of the diversity of Reddit moderators rather than being driven by a few active moderators.

Our work also raises interesting questions about how Reddit may react to moderators' protests, such as making their subreddits private or quitting. What if Reddit decides to hire commercial content moderators rather than spending time and resources addressing volunteer moderators' concerns? Given volunteer content moderators' close connection with communities and in-depth knowledge about community dynamics, it is unlikely for Reddit to replace volunteer content moderators altogether. However, some subreddits' moderators have expressed interest in asking Reddit to hire commercial content moderators to supplement their labor. A fruitful area of research would be exploring how to bring volunteer and commercial content moderators together to manage online communities.

Implications for the Public

By measuring online volunteer work's role in supporting businesses like Reddit, our work can better inform the public of ways to meaningfully shape the technology landscape. Currently, as technology companies that rely on volunteer labor gain more and more power in the technology realm, there exists evidence of their business development's negative impact on the public, e.g. monopolistic practices [**Kim and Luca2019, ?**] and lack of transparency in data use [**Sadowski, Viljoen, and Whittaker2021**]. By making explicit technology companies' dependence on volunteer labor, our work points to an opportunity for the public to collaborate with online volunteers to mitigate companies' influence over the technological ecosystem. Currently, online volunteers and, more broadly, members of the public supply valuable time, data, and knowledge to for-profit technologies such as social media platforms and rating systems but have little say over how these technologies are designed and developed. Our study highlights a potential direction that the public may take to mitigate this power imbalance. For example, the public may join online volunteers' in stopping their use of a technology or migrating to competing technologies—what Vincent et al. coined as "data leverage" [**Vincent et al.2021b**], to directly divert the crucial data and labor away from certain companies.

Implications for Policymakers

Finally, our work can assist policymakers in drafting regulations that can account for unpaid labor subsidies for for-profit technology companies. Economists have noted the sector's declining labor share and exacerbating income inequality [**?, ?, Posner and Weyl2018**]. Our method that estimates online volunteer work's value based on the work's comparable market rate could be adapted to different privatized online volunteer work, ranging from user-generated ratings to image labels. Previously, the monetary value of online volunteer work was difficult to estimate in part due to the complexity and opacity of the large, for-profit technologies it supports. For example, how much ad revenue volunteer moderators bring to Reddit when they remove a harmful post remains nebulous and may require sophisticated experiment design if at all possible. Our method is generalizable to other types of online volunteer work (see more details in Future Work) and could help policymakers understand how much value technology companies benefit from online volunteer work [**Au-Yeung2019, Commission2017**]. To more equitably distribute the profits of volunteer-powered technologies, just as tax assessors evaluate a property's market value to determine its owner's tax bill, policymakers could start assessing technology companies' "volunteer-dependence" tax based on the amount of online volunteer work calculated in a similar fashion to this paper.

Policymakers may also fund third-party "volunteer labor auditors" that conduct independent time studies to advocate for union members. Such entities could collect and analyze time logs from volunteers while preserving their privacy. Their findings would then assist volunteers in collective negotiation with companies that benefit from this labor.

Reflection on Data Collection

Our work is made possible due to the generosity of volunteer moderators who shared their mod logs and the openness of subreddits whose mod logs are made public. Given the sensitive and private nature of our data, we took special caution in our data collection, storage, and analysis. For those subreddits we recruited ourselves, we made our data collection script accessible to moderators so they could directly see what information we would anonymize and collect. In accordance with our IRB protocol, we cannot make mod logs we collected through our own recruitment public because mod logs contain extremely sensitive information about moderators' behaviors such as who removed what content. During recruitment, several moderators contacted us to confirm that only our research team would have access to their mod logs. Even if we anonymized all target links and account names, other information such as timestamps and action type may still risk moderators being identified. For the mod logs we collected from u/publicmodlogs, we cannot publish them because the mod logs are no longer publicly available. u/publicmodlogs only publishes the past three months' mod logs from its affiliated subreddits. By the time of this writing, all the mod logs we collected from u/publicmodlogs are no longer accessible online. We respect this setting to preserve the integrity of what moderators may have agreed to in adding u/publicmodlogs to their subreddits.

Future Work

Future work could extend our study in at least three directions. First, although our dataset is the most comprehensive dataset about moderation work on Reddit, our estimate is conservative. This estimate could be improved with more tracking of moderator behaviors. Moderator logs do not include time spent on untraced activities like responding to moderator mail, debating about moderation decisions on other platforms like Discord or Slack, and developing moderation bots. Additionally, our estimate only considers moderators of public Reddit communities because data about private communities' moderators is not accessible. Therefore, our estimate of moderator labor is a floor of the true amount of time Reddit's moderator population spends moderating on the site. Future work may enhance our estimate by analyzing moderator activities across tools and platforms from both public and private communities. Second, the hourly rate of commercial content moderation service from UpWork may underestimate moderation work's worth. Prior research on Amazon Mechanical Turk shows that monopsony power drives down crowd workers' wages [**Dube et al.2020**], which may also occur on UpWork to a lesser degree. Future work may collect more wage data to improve our estimate. Nonetheless, this approach can provide an important starting point about the value of online volunteer work. Third, our analytical method may be generalized to other types of online volunteer work to understand the amount of "labor subsidy" online volunteers supply to additional for-profit technology companies. Future work could explore the value of the labor hours online volunteers spend on writing reviews for products and services, providing implicit and explicit feedback for intelligent models' output, and producing answers on QA websites. For example, to assess the monetary value of the ratings volunteers provided on platforms such as Google Maps, one could first construct a reasonably representative sample of online volunteers and conduct a large-scale data collection to infer the hours they spent. Then researchers could model these volunteers' hours with publicly available user metrics such as years active and number of ratings written and use the model to extrapolate to the whole population for an estimate of the total volunteer hours. Finally, since providing ratings is a type of crowdsourcing task on Amazon Mechanical Turk, researchers could use the corresponding wage rate to estimate the monetary value of the estimated volunteer hours.

Conclusion

Using Reddit moderation as a case study, we estimate the monetary value of the online volunteer work completed by all Reddit volunteer moderators. This estimate may assist companies, volunteers, and policymakers in proactively upholding the volunteer-driven business model, the foundation of many successful technology companies and technological advancements. Our estimate is enabled by a novel method that projects online volunteer work's value based on equivalent, commercial services, and has the potential to be adapted for volunteer work beyond content moderation.

Chapter V: Out of Site - Opportunities for Collective Protests by Data Laborers

As protests by data producers have been increasing in both prominence and number, a growing number of activists are using the Internet to leverage the collective purchasing power of consumers. However, despite the attention given to recent online boycott campaigns, their success has been impeded by a number of significant challenges [**Friedman2002, G King2011**]. For instance, it can take a great deal of individual cognitive effort to conform to the campaign's goals, with individual participants having to remember a large, potentially changing blacklist of companies before every purchase [**Carrigan and Attalla2001**]. This can be even more difficult if, as is often the case, the target of a boycott owns a web of subsidiaries and a diverse array of brands. Additionally, it is often unclear to both organizers and participants how effective a boycott has been, with no way to track prevented or diverted purchases. This confusion and lack of visibility add to the already-significant sociotechnical obstacles associated with organizing any collective action campaign online, boycott or otherwise [**Klein, Smith, and John2004, Ling et al.2005, OLSON2009, Rashid et al.2006**]. To address these and other challenges to online boycotts, this paper contributes a new system called Out of Site. A prototype of a new class of applications that we call boycott-assisting technologies, Out of Site uses lightweight automation to eliminate some of the obstacles to successful online boycotts. Through this automation, Out of Site is also able to track the effectiveness of a boycott and share this information with all campaign organizers and participants. This allows online boycotts to much more easily benefit from social computing's large literature on incentivizing participation in online communities [**Cheng and Bernstein2014,**

Kraut et al.2012, Ling et al.2005, Rashid et al.2006, Zhu et al.2013]. Out of Site is implemented as a Chrome and Firefox web extension that allows users to join in any number of boycott campaigns. Once a user has joined a campaign, key campaign actions are automatically handled by the extension. For instance, when searching Google, Out of Site automatically hides or flags search results pointing to websites owned by a targeted company. Similarly, while shopping on Amazon, Out of Site hides or flags products manufactured by a targeted company. We designed Out of Site such that campaigns are easy to design by non-technical users, democratizing access to this new type of automation-assisted boycott.

Figure 13 shows a screenshot from a user who has joined the Out of Site campaign we implemented for GrabYourWallet. GrabYourWallet is a boycott community that seeks to exert economic leverage on U.S. President Donald Trump, particularly in response to the Access Hollywood tape scandal [**Wom**] and the subsequent sexual assault and harassment allegations [**Grab**]. In Figure 13's screenshot, which captures a query for "Hobby Lobby" (a company targeted by GrabYourWallet), one can see that Out of Site filtered out links to the company's website and removed the locations of its physical stores from the included map. To understand how users interact with boycott-assisting technologies like Out of Site, we deployed Out of Site with members of two existing boycott communities: GrabYourWallet and the animal rights campaign Stop Animal Testing. Our deployment consisted of two phases: an 18-day study of our first version of Out of Site and a follow-up four-week study of an improved second version of the system. During the first study, we gathered log, survey, and interview data from 42 participants from these two communities. Following this first study, our research team integrated participants' feedback and

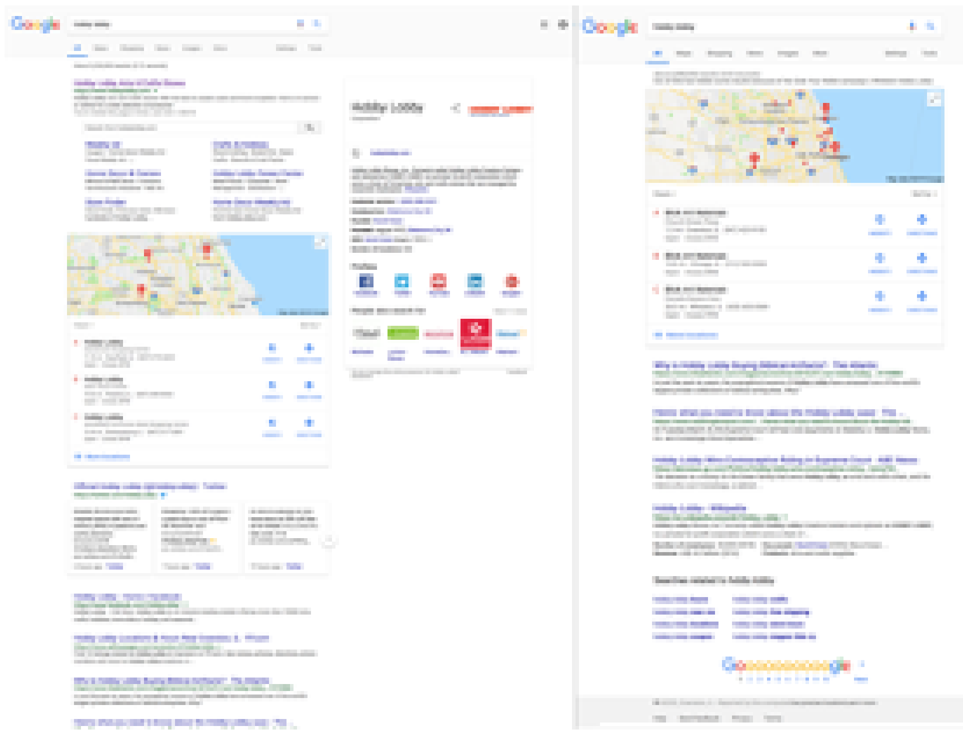


Figure 13. Left: A screenshot of unaltered Google search results for "hobby lobby". Right: The same query with Out of Site running the GrabYourWallet campaign (in "High" mode; see below). Note that almost all of the search results have changed, with the top results on the right all going to Hobby Lobby competitors. The links that made it through Out of Site's filters are discussed below.

additional design implications from the literature to improve Out of Site. We then deployed a second version of the system and collected more log data from the 26 first-study participants who kept the extension installed and did the same for 19 new participants that we recruited. The results of our deployments show that Out of Site had a meaningful impact on our participants' web experiences, with hundreds of webpages affected. We saw evidence of participants' traffic being redirected to the websites of competitors of targeted companies. More generally, participants expressed excitement about the idea of boycott-assisting technologies, although they revealed diverse preferences regarding the specific implementations of key components of these technologies (e.g. whether targeted content should be hidden or simply flagged). Interestingly, however, we also observed (and were able to interview our participants about) occasional intentional non-conformance behaviors. Specifically, some participants chose to whitelist certain targets, such as Amazon, or found workarounds to access targeted companies' websites. While the primary contribution of this paper is the Out of Site system and its approach to online boycotts, this paper also makes a secondary contribution in the design approach we used to implement Out of Site. Computing researchers have recently come under increasing criticism for considering only the positive impacts of a paper's contribution in cases when the literature and current events make it clear that the contribution will have predictable negative impacts (e.g. [**AIG2018, Bigham2018, Frauenberger et al.2017, Griffith2017a, Parikh2018, Vardi2018**]). Using a traditional design approach for Out of Site would have made this project quite liable to this critique. For instance, a system that empowers just any boycott could easily enable a powerful boycott against African American-led businesses or against businesses owned by religious minorities. Out of Site could also be used to severely bolster a user's filter bubble [**Van Alstyne and Brynjolfsson2005**], isolating the user from any new information that might change the user's opinions about a boycott target.

To address these and other predictable anti-social uses of Out of Site, we developed a very straightforward design approach that we call heuristic preventative design (HPD). The goal of this approach is to, with minimal burden on the researcher or developer, meaningfully mitigate the negative impacts of a given system or contribution. HPD involves using the literature and current events to identify a blacklist of uses of a given system and, early in the design process, making design decisions to make those uses much more difficult. We describe how we used HPD to put obstacles in the way of the above uses of Out of Site (and others) while at the same time ensuring that Out of Site supports a diverse array of values and perspectives. We also highlight how the HPD process is sufficiently simple – it does not require specialized knowledge or extensive training – that it should be immediately accessible to a broad range of the computing community.

Below, we first present work from several disciplines that helped to motivate the idea and design of Out of Site. Next, we walk the reader through our design process. Finally, we present the findings from our deployments and discuss the implications of these findings.

Related Work

In this section, we discuss the prior work that helped to motivate Out of Site and its design. This work emerges both from social computing as well as from several other domains, including consumer studies and political science.

Consumer Boycotts

Consumers use boycotts both to serve economic objectives, such as lowering prices, and to fulfill political and societal objectives, such as supporting fair trade [**Atkinson2012**]. In political science, boycotting is understood as a form of political consumption, i.e. people making purchasing decisions for political or ethical reasons [**Micheletti, Follesdal, and Stolle**]. Multiple studies provide evidence showing that political consumption has become a prominent form of political participation and civic engagement [**Andorfer and Liebe2012, Barnett et al.,**

Stolle, Hooghe, and Micheletti2005]. Indeed, political consumption research suggests that there is a large population of potential users for systems like Out of Site. For instance, in a recent U.S. survey , more than 28 percent of respondents reported engaging in at least one form of political consumption [**Newman and Bartels2011**]. Similar studies in European countries have also observed substantial interest in political consumption, e.g. 35However, despite the popularity of boycotts among consumers, many barriers to successful boycotts still exist. To understand these barriers, it is useful to consider the consumer behavior literature on boycotts. Specifically, this literature has identified two types of potential boycott participants: (1) the “caring and ethical” type and (2) the “confused and uncertain” type [**Carrigan and Attalla2001**]. “Confused and uncertain” consumers are those who attempt to join a boycott but are often overwhelmed with messages they perceive to be contradictory and do not receive enough guidance regarding specific actions [**Klein, Smith, and John2004**]. Out of Site was designed to reduce the cognitive burden on “confused and uncertain” boycott participants by streamlining and automating boycott actions (see Section 3). Out of Site also mitigates barriers to successful boycotts for “caring and ethical” consumers. These consumers are known to struggle to integrate new information about boycott targets [**Boulstridge and Carrigan2000**] and, as described below, Out of Site performs this integration automatically. Aside from facilitating boycotts, Out of Site can also be understood as a system that supports a related political consumption strategy: the technique known in the political consumption literature as the “buycott” [**Hawkins2010**]. While boycott participants avoid purchasing goods, buycotts involve consumers purposefully purchasing goods from desired businesses, e.g. purchasing fair trade products [**Andorfer and Liebe2012**]. While boycotting adopts a conflict-oriented strategy to punish bad companies, buycotting is centered around a cooperation-oriented strategy to reward good companies [**Friedman2002**]. Out of Site improves consumers' exposure to non-target companies by deprioritizing boycott targets in search results, e.g. the relationship between Blick Art and Hobby Lobby in Figure 1. Through this approach, Out of Site combines some of the benefits of buycotts with those of boycotts. Many modern political consumption campaigns are initiated by individuals online and later attract collective attention [**Bennett and Segerberg2012**]. One successful example is the deleteUber hashtag. This boycott started when a Twitter user called people's attention to an Uber promotion that took place in the context of a taxi driver strike related to changes in U.S. immigration policy [**Isaac2018**]. The hashtag became trending in the U.S., which led Uber to make a public apology and set aside funds to support Uber drivers who were affected by the immigration policy [**Sta2017**]. The success of the deleteUber campaign, however, is an outlier. Researchers have studied multiple cases of individual-initiated online boycotts and found that these grassroots boycotts rarely become effective in terms of posing economic harm to the targeted companies. Out of Site was motivated in part to improve this success record, and to do so, this research suggests that adopting approaches from the collective action literature will be necessary. We discuss how the collective action literature helped to motivate Out of Site's design below in the sub-section that follows. While Out of Site is by far the most advanced boycott-assisting technology of which we are aware, there are some basic tools available whose design helped to inspire Out of Site. In particular, a number of boycott campaigns have searchable databases of targeted companies and products on their websites. Out of Site effectively integrates these databases directly into participants' web browsing experiences and automates key boycott actions based on these databases. Similarly, there are a few lightweight applications that provide different interfaces to access these types of databases. One interesting such application is Buycott [**Buy**] (not to be confused with the formal term from the political consumption literature). Buycott is a mobile app that uses bar codes to query targeted product and company databases. However, Buycott is "read-only"; it does not attempt to automate any boycotting actions as in the key features of Out of Site.

Collective Action

Social computing researchers have examined how well-organized online collective action has led to policy changes [**Matias2016d**], social awareness [**Dimond et al.2013**], successful crisis response [**Starbird2013**], and other positive outcomes. While not directly related to boycotts, the lessons learned in the social computing collective action literature has many applications to the boycotting context. In particular, in offline boycott campaigns and existing online boycott campaigns, the number of participants and the aggregated economic impact are often not visible. As such, boycott participants have no way to know whether a campaign is gaining traction or has made a difference [**Klein, Smith, and John2004**]. The collective action literature in social computing (and other fields) suggests that increasing the visibility of collective progress incentivizes sustainable participation and can potentially lead to larger impacts (e.g.

[**Ling et al.2005, Shaw et al.2014, Vasi and Macy2002**]). To implement this design implication, Out of Site measures and provides real-time feedback to users about the collective progress of their boycott campaign(s). Specifically, as described in the "Design" section, the extension keeps track of how many participants have joined a boycott and the aggregate impact on participants' web experiences. There is, however, one well-known caveat to the observed benefits of group feedback: at the beginning of a collective effort, seeing that only a few people are participating can provide a strong individual disincentive to participate [**Cheng and Bernstein2014, Granovetter1978**]. This creates a well-known "chicken and egg" problem for online communities [**Kraut et al.2012**]. To address this problem, many design strategies have been proposed, ranging from paying professional users to attracting endorsements from prominent individuals (other strategies include creating scarce resources for early adopters, deploying bots; see Resnick et al. [**Kraut et al.2012**] for an overview). To address this problem, we implemented a version of the former strategy (professional users) by including boycott statistics from our pilot testers prior to the first phase of our formal deployment. In addition to practical online community design, this approach also had the benefit of ecologically validity: boycott campaigns in Out of Site will almost always be first used by individuals in the organizations that design them (see below) before they are successful with the general public. An issue related to the "chicken and egg" problem is the "free rider problem" [**OLSON2009**], which broadly describes when an individual gains the benefits of a group effort without making proportional individual contributions. Prior work has shown that making visible individual contributions to group progress can mitigate (and even reverse) this problem in online collective action contexts (e.g.

[**Ling et al.2005, Rashid et al.2006**]). Based on this literature, we designed Out of Site to prominently highlight individual boycott contributions as well as displaying group progress. Another well-known challenge in collective action is coordination, especially when collective action campaigns reach the large scales they often need to have a desired effect [**Oliver, Marwell, and Teixeira1985, OLSON2009**]. Out of Site has a number of features that are designed to help mitigate this challenge. For instance, the list of campaign targets is maintained server-side, meaning that changes to targets by organizers will be quickly propagated to campaign participants. Finally, when considering collective action in online contexts, one type of technology that cannot be overlooked is social media. Extensive literature has demonstrated how social media can help catalyze successful collective action campaigns (e.g. in crowdfunding [**Lu et al.2014**], social movements [**Bozarth and Budak2017, De Choudhury et al.2016**], crisis response [**Starbird2013**]). Additionally, political consumption studies suggest that social media users are more likely to participate in boycotts and buycotts [**Becker and Copeland2016, de Zúñiga, Copeland, and Bimber2014**]. To leverage the power of social media, Out of Site contains a social sharing feature (described in more detail below) that enables users to directly engage with their social networks around their participation in Out of Site campaigns.

Browser Extension Research

Our decision to implement Out of Site as a browser extension was motivated in part by several recent projects that have highlighted how browser extensions can be used to alter the black-box behavior of private technologies (e.g. search algorithms) and can assist with online activism more generally. For instance, in an effort to understand how much Google depends on Wikipedia links to satisfy user information needs, McMahon et al. [**McMahon, Johnson, and Hecht2017b**] built a browser extension that silently removed Wikipedia links from Google's search results. They deployed the extension in a small study, finding that Google search performance dropped substantially in many cases when Google could not surface Wikipedia links. This research highlighted for us the power of altering the search experience and helped to motivate the related functionality in Out of Site, although Out of Site allows campaigns to customize the search experience in a much more extensive fashion. Browser extensions' advantages are not limited to altering search experiences; other studies have implemented browser extensions to address other power imbalances in social computing systems [**Howe and Nissenbaum2017a, Irani and Silberman2013b, Munson, Lee, and Resnick2013**]. Turkopticon [**Irani and Silberman2013b**], built on top of Amazon Mechanical Turk, is an activist system that helps workers to publicize and evaluate requesters, a function that AMT does not natively afford. In a similar vein, Howe and Nissenbaum built an extension that allows users to directly act against online advertisers by obfuscating clicks on ads [**Howe and Nissenbaum2017a**]. The extension, AdNauseam, extends the idea of ad blockers and simulates random clicks on the blocked ads to confuse trackers. These systems demonstrate the potential of browser extensions in empowering users to contest powerful entities, an idea that is central to boycotts. Our decision to use a browser extension-based approach was most directly motivated by GrabYourWallet's launch of its own very lightweight browser extension approximately one year ago. The GrabYourWallet extension, which was a popular request from GrabYourWallet participants [**Graa**], has a single, simple function: it alerts the user when the user has gone to a website of a company that is on the GrabYourWallet target list. Out of Site can be viewed as a generalizable, much more powerful version of the GrabYourWallet extension. While Out of Site also can notify users when they attempt to visit a targeted website (we intentionally subsumed the functionality of the GrabYourWallet extension), it also does the difficult and critical work of customizing Google and Amazon search results pages (e.g. to benefit the millions of people who purchase through Amazon rather than going directly to company websites), works on a product-by-product basis rather than blocking entire websites, and has the extensive list of additional functionality outlined in the Design section below. Moreover, Out of Site additionally adopts design implications from the literature mentioned above (e.g. political consumption, social computing) to better facilitate collective action from interested participants. Our hope was that by subsuming the functionality of the GrabYourWallet extension while building a qualitatively more powerful system, we could build on the enthusiasm surrounding the extension to attract participation.

Design

The primary contribution of this paper is a system and, as such, the primary methodological challenge of the paper was the design and implementation of the system. In this section, we outline our design process and provide details about how our implementation was motivated by related work. We also discuss how we developed and utilized heuristic preventative design (HPD), our lightweight design approach that integrates the mitigation of some negative uses of a technology directly into the design process. Finally, as Out of Site is a system that has been publicly deployed and thus has required frequent iterative design improvements, we also highlight the many more minor changes we have made to the system directly in response to user feedback.

Design Objectives

After studying the online boycott ecosystem and the existing literature, we developed three design objectives for Out of Site. The first two objectives directly address major problems in existing approaches to online boycotts, as discussed in the Related Work section above. The third objective emerged out of increasing calls for the computing community to pay more attention to the potential negative uses of the technologies it develops (e.g. [**AIG2018, Frauenberger et al.2017, Hecht et al.2018, Parikh2018, Vardi2018**]). More specifically, our three objectives are as follows: 1. Conformance to Boycott's Goals: Help people individually conform to the goals of a boycott by offloading to an automated system the logic of figuring out which items are boycotted and acting on these items (as discussed in Section 2.1 above). 2. Collective Action: Help people both act collectively and track the progress of a boycott (as discussed in Section 2.2 above) 3. Avoiding Negative Impacts: Achieve goals 1 and 2 while minimizing significant and predictable negative impacts. Below, we organize our discussion of our design decisions using the framework provided by these three objectives. We then close by discussing iterative improvements to Out of Site that were motivated by user feedback rather than the literature.

Design Approach

Conformance to Boycott's Goals: As discussed above, a major challenge faced by boycott campaigns is the tremendous cognitive effort associated with successfully participating in a campaign [**Carrigan and Attalla2001, Newholm and Shaw2007**]. This is a problem regardless of whether boycotters fall in the "confused and uncertain" category or the "caring and ethical" category (see Section 2.1). Boycott participants in the former category struggle to keep track of current targets and to recollect all of these targets' various subsidiaries and brands. In this sense, Out of Site needs to onboard this cognitive effort by identifying targets accurately and correctly in real time, and then help these participants take action on this knowledge. Out of Site also benefits consumers in the latter category, who face challenges in integrating new information at the point of every purchase [**Klein, Smith, and John2004, Newholm2005**]. As such, Out of Site needs to closely and timely integrate any target changes into the system. In standard human-computer interaction terms, these issues can be understood broadly as placing excessive burden on users' recall capacity, and we know that computing technology excels at reducing recall burden [**Dingler2016**]. As such, this major challenge to successful boycotts seemed particularly well-suited to address using computing technology. The decision to implement Out of Site as a browser extension emerged directly from this design goal, along with the prior work in browser extensions discussed above. Browser extensions can observe and change virtually any web experience (e.g. [**Kim, Hullman, and Agrawala2016, McMahon, Johnson, and Hecht2017b, Munson, Lee, and Resnick2013**]), and thus can easily monitor, for instance, when a user is on a shopping website or is using Google to search for a product. They can then intervene accordingly, requiring no user recall. This event-driven behavior, however, only assists the user in remembering to act on the boycott when they are engaging in commercial behavior. To address the recall issues associated with keeping an up-to-date copy of all brands and subsidiaries owned by all targeted companies in a given boycott, we implemented Out of Site with a server-side backend that keeps track of this information and can be updated at any time by campaign organizers. The backend is implemented as a straightforward set of keywords and web domains. These keywords and domains are currently specified manually by the campaign organizer (although we are planning a system that would automatically generate these with minimal human input). For instance, one of GrabYourWallet's target companies is the G-III Apparel Company. For this company, the set of keywords include "Calvin Klein" and "Tommy Hilfiger" and the set of domains include "tommy.com" and "calvinklein.us".

Out of Site uses the lists of keywords and domains as the primary inputs to its six intervention types: filter, rerank, gray-out, call-for-action, block, and redirect. The filter intervention type removes DOM elements whose content contains any of the keywords in a campaign dataset or, in the Google case, contains links to any of the domains (filter is the intervention type featured in Figure 1). Filter requires non-trivial custom engineering to provide a good user experience on a given website, and as such, support is limited to Amazon and Google (although expanding it to sites like Wal-Mart, eBay, and Bing would simply require following the same development process we used for Amazon and Google). The main engineering challenge comes from special cases, e.g. Google's Local Search results; these require unique treatment relative to other types of DOM elements.

The rerank intervention type works on search results output by Amazon and allows campaign organizers to place a targeted search element at the bottom of a ranked list on a webpage rather than filtering it out entirely. Similar to filter, this is an intervention type that can be extended to other websites with search engines, but it does require some custom engineering. The gray-out intervention type works quite similarly to filter, but instead of removing the targeted DOM element, it places a semi-transparent box over the element. This intervention type also shows a message indicating that the element contains a link or product that is targeted by a given campaign. Our implementation of gray-out is supported for both Amazon and Google. The call-to-action intervention type builds on top of gray-out, but instead of modifying the DOM element, the extension injects a campaign's call-to-action (e.g. sending an email to the company's PR team) around the element. This intervention aims to help boycotts that encourage participants to be vocal and directly communicate with targeted companies. For instance, for the GrabYourWallet campaign, when this intervention type is used on Google, the following moderately-sized text appears above a targeted DOM element: "Company is targeted by the campaign GrabYourWallet." It then asks the user to "consider contacting the company" to express dissatisfaction with company behavior. Call-to-action also provides a link to an e-mail pre-addressed to the company's e-mail address (as suggested by the GrabYourWallet campaign). The user only needs to press "send" in their e-mail client to complete the action.

The block intervention type is the most straightforward: it simply blocks users from going to a specific targeted domain, providing a message that the domain is inaccessible due to the Out of Site campaign (and provides the campaign's call-to-action). The redirect intervention type is similar to block, but instead of blocking the site, it redirects the user to the boycott campaign homepage after showing a message. In both these intervention types, users' visits to a targeted domain are interrupted only once within an hour. After the interruption, users can access the targeted domain without any interruption for one hour if they choose (see below).

Collective Action: As described above in Related Work, Out of Site was designed from the ground up to incorporate several key design implications from the social computing literature on collective action. Here, we provide more details about our approach. Communicating group progress and individual contributions are crucial for collective action success (e.g. [**Ling et al.2005, Rashid et al.2006, Shaw et al.2014, Vasi and Macy2002**]). Out of Site provides two types of feedback about group progress: the number of current campaign participants and the potential total impact of the campaign. Calculating the number of participants is a simple server-side measurement, and this information is displayed in the main drop-down menu of Out of Site (see Figure 4). Calculating the impact is more difficult and varies across types of campaigns. For instance, GrabYourWallet calls for participants to eschew shopping at the targeted companies. As such, for GrabYourWallet (and campaigns with similar goals), the extension highlights in its main drop-down interface how many visits to targeted websites have been blocked and how many search results to targeted websites have been altered (Figure 4). The goal of Stop Animal Testing is to avoid buying specific products. As such, for this campaign (and campaigns with similar goals), the extension displays the total number of products that have been hidden on Amazon (Figure 4). Because prior work has shown that highlighting individual contributions to group progress will increase participation in group activities (e.g. [**Ling et al.2005, Rashid et al.2006**]), Out of Site also displays each user's contributions to group progress for both types of campaigns. A new social sharing feature was implemented for the second phase of our deployment. This feature was motivated by the literature on social media and online collective action (e.g. [4,72)], and was also developed in response to user feedback that encouraged leveraging social media. By clicking on a button in the Out of Site interface, a pre-filled template message is presented that users can edit or post as-is to Twitter or to a user-selected group of e-mail addresses (implementing support for other social networks would be relatively straightforward). Determining the best design for our template message was the subject of some literature review. In both collective action and personal informatics, researchers found that sharing active status reports are better received by audiences [**Epstein et al.2015, Flores-Saviaga, Keegan, and Savage2018**]. Relatedly, in existing studies about collective action on social media, hashtags have been identified as important in building and facilitating online communities [**Bozarth and Budak2017, De Choudhury et al.2016**]. Wikipedia researchers have also found that messages that highlight specific actions more successfully increased contributors' attention to a desired task [**Zhu et al.2013**]. As such, our template message includes (1) personal contributions to a campaign, (2) appropriate hashtags, and (3) a specific call for followers to join the Out of Site boycott campaign. For example, a GrabYourWallet user's message might appear as follows: e.g. "I boycotted 47 websites to support GrabYourWallet using Out of Site (a Chrome extension). Join me now: http://bit.ly/2IkmxCq. Read about the campaign: http://grabyourwallet.org."

Avoiding Negative Impacts: Our review of the literature and current events made it clear to us early in the design process that there are a number of obvious ways that Out of Site could be used by nefarious or well-intentioned actors to create negative impacts. These predictable negative impacts include the creation of severe filter bubbles and the use of Out of Site's boycott-assisting power by hate groups or extremists, e.g. neo-Nazis wanting to boycott all companies with Jewish and African-American CEOs.

The default approach in computing research – as was recently highlighted in a proposal from the ACM Future of Computing Academy [**Hecht et al.2018**] – is to consider the prevention of negative impacts like those above to be out of scope for a technical paper. That is, under the computing research status quo, the responsibility of the computing researcher who develops a new technology is the development of the technology, not mitigating the predictable negative impacts of that technology. For instance, those who develop generative models for audio and video are not expected to find ways to mitigate the disruptions to democracy that may be caused by these models [**Dee2018**].

With Out of Site, we wanted to try to do something different. Specifically, we sought to find a tractable design approach by which we could treat preventing predictable negative impacts as a first-order design concern rather than as an afterthought. Because of the urgent need for such an approach in a wide variety of computing research projects [**AIG2018, Cummings2006, Hecht et al.2018, Vardi2018**], we also sought to identify a generalizable process rather than one specific to this project.

As Hecht and colleagues and Parikh note [**Hecht et al.2018, Parikh2018**], one reason why computing researchers fail to engage with the negative impacts of their research is that the effort associated with making a technical contribution in computing is already very extensive. Computing researchers have also argued (e.g. [**Bigham2018, Cummings2006, Parikh2018**]) that computer scientists are usually not trained in the diverse array of social science methods and critical theory necessary for a more complex engagement with these negative impacts (e.g. [**Frauenberger, Rauhala, and Fitzpatrick2017, Hayes2014, Muller2003b**]). For this reason, our goal was to develop a low-cost, very straightforward design approach that can address a non-trivial subset of negative impacts. While quite incomplete, such a design approach would still be a substantial improvement over current practice and would be much more easily adoptable by a wide range of computing researchers. This is similar to the logic that motivates the extremely well-known user interface evaluation technique called heuristic evaluation [**Nielsen and Molich1990**]. Heuristic evaluation gives designers a portion of the benefits of costly user studies, but is much quicker, simpler, cheaper, and more straightforward. This makes it accessible to a much wider group of computing professionals and much wider set of computing projects.

The design approach we developed we call heuristic preventative design (HPD), and it was inspired both by the pragmatism of heuristic evaluation and a debate that has occurred in the literature on value-sensitive design (e.g. [**Cummings2006, Friedman, Howe, and Felten2002, Howe and Nissenbaum2017a**]). With respect to the latter, an early view of value-sensitive design was based on an assumption of universal values between designers and nearly all users, with designers encouraged to incorporate these values into systems [**Friedman1997, Friedman and Kahn2003**]. Le Dantec et. al. later advocated for a different understanding of value-sensitive design, one that problematized the notion of most values being universal and instead advocating that we design our systems to conform to the values of their users [**Le Dantec, Poole, and Wyche2009**]. HPD leans heavily on the latter interpretation, but includes a critical, heuristic evaluation-like component of the former.

Like more recent interpretations of value-sensitive design, HPD asks designers to support and enable a wide variety of user values. For Out of Site, this means making it possible for a diverse set of boycott organizers with a diverse set of values to gain the benefits of the system (e.g. politically progressive or conservative boycotts). The heuristic prevention emerges in the form a specific, small set of blacklisted uses of the system that is developed by the research team. This blacklist amounts to a heuristic for the negative impacts of the system. The definition of "negative" here comes from the research team, as per the traditional understanding of value-sensitive design. The team then modifies the system design to specifically prevent the blacklisted uses.

A key component of HPD is the method one uses to develop the set of blacklisted uses. As per the pragmatism of heuristic evaluation, we suggest (and took) the following lightweight approach: following the completion of the literature review process, the research team had a meeting to brainstorm a list of potential misuses of Out of Site that could be predicted by the literature review. Given the prominence of the negative impacts of computing systems in current events (including those of social computing systems), the team also engaged with relevant recent news stories as well. Finally, motivated by the even more pragmatic approach of guerrilla usability testing [**Nielsen1994a**], we also discussed the topic of potential negative uses of Out of Site with several colleagues. Within a very short period of time, we had developed a list of four potential negative uses of our system that could be easily predicted from prior work and current events. These uses and how we sought to prevent them are discussed in more detail below. We also discuss below how individual project blacklists could be aggregated into a global list that would further reduce barriers to HPD application.

As discuss earlier, HPD partly relies on an older understanding of value-sensitive design in which there is some universal value set that researchers can use to make decisions about what is "use" and what is "misuse", who is a "nefarious" actor and who is a "virtuous" actor, what is a "negative" impact and what is a "positive" impact. While it is an (intentionally) imperfect assumption, the assumption of HPD is that there is a large set of uses of technology that would be considered misuse by enormous segments of the population. Examples prominent in current events include disruptions of democracy, technology addiction, and undetected and severe violations of privacy. While some may object to designing technology to prevent these broad uses, most would not. The same is true, for instance, regarding the neo-Nazi uses of Out of Site. In other words, HPD favors pragmatism over perfection in this respect. HPD is not the only approach to utilize this large-majoritarian approach; ACM leveraged a similar approach when it developed its new ethics guidelines [**ACM**].

Ultimately, however, if a researcher or user disagrees with a blacklist choice, she or he can develop a similar technology that supports these uses. We expect that if there is a reasonable argument for the mistaken inclusion of a use on a blacklist, this would be a good motivation for a new and successful research project. However, if no such argument exists, time, money, and other factors (e.g. social costs as discussed by the FCA proposal [**Hecht et al.2018**]) will act as significant barriers to this type of activity.

These barriers could perhaps be strengthened even further through the use of intellectual property law – e.g. patents – which may give researchers and designers the legal power to exclude others from the use or reimplementation an invention for a period of time. Under current practices, these barriers to blacklisted uses do not exist, and, as a result, computing researchers are effectively subsidizing these uses.

We emphasize that HPD with its small set of blacklisted uses is not intended to replace a more rigorous consideration of negative impacts that engages much more seriously with the social sciences and with critical theory. Instead, HPD is intended merely to replace the default status quo: doing nothing. Our expectation is that the HPD approach can be a stepping stone to more advanced, more complete approaches that require more training. For instance, action research [**Hayes2014**], participatory design [**Muller2003b**], and ethos building [**Frauenberger, Rauhala, and Fitzpatrick2017**] may be appropriate more advanced approaches for projects in which HPD is used. However, as can be seen with our implementation of HPD for Out of Site (described immediately below), even this stepping stone can result in real and concrete changes to a technology.

Implementation of Heuristic Preventative Design (HPD) in Out of Site: As discussed above, HPD can be broken down into three key steps:

(1) Ensuring that a system is adaptable to a diverse set of user values. (2) Building a blacklist of system uses that is motivated by the literature and current events. (3) Designing of the system to prevent the uses on the blacklist.

Our approach to implementing the first step was to ensure that boycott campaigns are very easy to create and highly adaptable to boycott organizer preferences. Organizers can choose arbitrary keywords and domains and can decide what happens to DOM elements and sites that match those keywords and domains on a customizable basis. Indeed, our boycott campaigns are implemented as simple JSON objects that describe the various intervention types that are desired and for which websites. No campaign-specific code is included in the extension itself. For instance, an organizer seeking to boycott Vista Outdoors (whose brands were recently dropped by REI because they also own a gun manufacturer) could simply enumerate the keywords associated with the campaign (e.g. "camelbak" and "giro", two brands owned by Vista) and the domains targeted by the campaign (e.g. "camelbak.com"). They would then select the interventions they wanted, e.g. on Google, use filter; on Amazon, use re-rank. Campaign organizers can also define the meaning of the "strength" levels (see below). All of these settings could easily be enumerated using a simple web wizard, for which we have built a prototype. However, we stopped development of that prototype for the reasons described below and we designed our two campaigns for our deployments using the raw JSON. The process we used to develop the blacklist for Out of Site was exactly the very lightweight process we described above. Following the completion of this process, we had developed a list of four blacklisted uses. We describe these uses in detail next. We couple our description of each blacklisted use with our implementation of the third step in HPD (altering the design of the system for that specific blacklisted use). Blacklisted Use 1 – Enhanced Filter Bubble: While there is some debate in the literature about where and when filter bubbles exist (e.g. [**Nguyen et al.2014**]), most scholars agree on the negative impacts of filter bubbles that do exist (e.g. [**Dillahunt, Brooks, and Gulati2015, Pariser2012, Resnick et al.2013, Van Alstyne and Brynjolfsson2005**]). This is problematic for Out of Site, as one could easily leverage Out of Site as a way to block out all undesirable information about a topic, instantly inducing a sort of "filter bubble on steroids". For instance, an Out of Site a campaign could easily be built to remove all information from search results about oil companies, including removing important news stories about these companies. Similarly, one could use the infrastructure of Out of Site to remove all information from search results about a political party one does not support (e.g. a supporter of the U.S. Democratic party could block out all search results that mentioned a member of the Republican party in a campaign with the appropriate keywords). We made specific design choices to prevent this "filter bubble on steroids" use of Out of Site. In particular, the simplest way to implement our Google SERP filtering would certainly have been to simply treat all search results in the same way, filtering out those that include targeted domains and keywords and allowing those that do not to surface to the user. However, to prevent this blacklisted use of Out of Site, we treat search results in a more nuanced fashion and Out of Site only supports the targeting of certain types of search results. In particular, Out of Site only targets search results when those results can be assumed to have a commercial intent. For instance, we do not target Wikipedia content, news stories, or any elements that are part of Google's news carousel. This means that a user who is boycotting oil companies would not see the websites for oil companies, but would see news related to the oil companies. This is particularly important in a boycotting context: news may emerge about a boycott target that might change a participant's view of the target. This news could take much longer to reach the participant without this adaptation. Another adaptation we made to prevent this use is the introduction of boycott "strength" levels that are customizable by participants. Out of Site allows campaign organizers to set three different levels of intervention types, and users can switch between these levels. The "High" strength level is intended to be the most invasive configuration (e.g. frequently using filter). "Medium" and "Low" are designed to move towards less invasive intervention types like re-rank and gray-out. The strength level adaptation allows users greater exposure to targeted information if they want it, e.g. rather than search results disappearing, they would have a call-to-action around them. Below, we see that some users in our deployment moved Out of Site from the default "High" setting to the more moderate "Medium" setting. Blacklisted Use 2 – Use by Nefarious Groups: Our initial plan for Out of Site was to develop an easy-to-use wizard to help any person create a campaign and to support the distributed dissemination of campaigns between users. However, given recent studies on hate speech [**Hine et al.2017**], misinformation [34], and trolling [**Flores-Saviaga, Keegan, and Savage2018**] (and related topics), it was clear that such an approach could be co-opted by hate groups to implement campaigns with clear negative impacts, e.g. the campaign mentioned above targeting Jews and African-Americans, and similar campaigns targeting companies owned by women. To make this use much more difficult, Out of Site now requires all campaigns to go through server-side activation. If Out of Site were to become popular, this would allow a gatekeeper to use a public policy to determine which campaigns would be supported. When instrumenting this server-side-only approach, we identified several concrete restrictions that could be included in this policy, e.g. (1) the organizer cannot be a hate group as defined by the Southern Poverty Law Center [**Gro**], (2) the campaigns can in no way selected targets defined by protected classes as defined by U.S law (supplemented by sexual preference and gender identity) [**Typ**], and (3) the extension cannot be used by a state actor. We note that such a gatekeeping approach is not novel to Out of Site. For instance, Apple's App Store takes a similar approach, in part for similar reasons [**App**]. Blacklisted Use 3 – Excessive Removal of Autonomy from Users: One controversial impact of computing systems that has attracted substantial attention is that they these systems are taking autonomy away from humans, especially in informational contexts. Ample research has shown how search engines, Facebook, and Newsfeed have impacted users' decision making (e.g. [**Epstein et al.2017, Kramer, Guillory, and Hancock2014, Lokot and Diakopoulos2016**]). Out of Site in the most extreme cases could exacerbate this concern. Our strength levels were designed as a protection against this concern. Additionally, as the filter intervention type directly removes content from users' web experiences, we insert cues in users' webpages to indicate some content has been hidden. Such cues include a short sentence stating "Out of Site has hidden some results because of the campaign name campaign" and the inclusion of the number of hidden items above the extension's icon in browser (see the top of Figure 4). Additionally, we allow users to whitelist individual targets with small in-context cues displayed on affected webpages such as "Whitelist company name — Whitelist another company name". We also allowed users to whitelist boycott targets by using the extension's detailed settings. Blacklisted Use 4 – Causing Undue Harm: In Out of Site, the call-to-action intervention type provides easy-to-use instructions for actions users can take to support a boycott, e.g. sending an e-mail to a boycott target. Such an affordance allows boycott participants to be vocal about their opinions, but if exploited, may cause excessive harm to targets (e.g. SPAM). In our implementation, we set a maximum number of daily call-for-actions that users could execute. Additionally, some of the design implications that emerge from our user study raise critical issues related to undue harm, and we highlight these issues in the Discussion section.

Iterative Design

Aside from the updates of Out of Site mentioned above, we have made several additional improvements to the extension in response to user feedback and usage data. Below we describe a few of the more significant of these improvements (all changes listed were made in time for the second deployment phase). Treatment of Third-Party Commercial Content: Initially, we considered information about companies on third-party commercial platforms to be non-commercial content, e.g. we did not filter or gray-out Yelp reviews. After all, many of the Yelp reviews could be poor, and regardless, Yelp reviews may might provide new, third-party information to consumers (as is the case for news articles). However, we noticed in users' search results that the costs to boycott campaigns of this approach almost certainly outweighed the benefits. For instance, in Figure 1, in an earlier version of Out of Site, a relatively positive Yelp review of Hobby Lobby appeared prominently in the post-filter results. As such, Out of Site now considers all prominent third-party commercial platforms to be commercial content and treats them accordingly (e.g. Yelp.com, coupon websites, links to app stores). Campaign Contribution Metrics: Originally, only when Out of Site was set to "High" were users' contributions to campaigns counted towards campaign progress metrics. However, users gave us feedback that when they use "Medium" setting, Out of Site still helps them to avoid the target, and thus "Medium"-level interventions should be counted as contributions. Based on this feedback, we updated this counting mechanism. Simplifying Contribution Metrics: Some users in our first deployment phase expressed confusion about the granular campaign and individual progress metrics that the first version of Out of Site provided. To avoid this confusion, Out of Site now presents the simplified metrics shown in Figure 4.

Installation Wizard: Our interviews following the first deployment phase made it clear that many users were not aware of some of the features in Out of Site, nor were they aware that they could opt-out of any of the installed campaigns. As such, for the second deployment phase, we implemented an installation wizard that had users opt in to campaigns and, for each campaign, choose the “High”, “Medium”, or “Low” setting. This change also had the added benefit of allowing more users to be exposed to the Amazon.com features in the Stop Animal Testing campaign (see below).

To better understand how people interact with Out of Site and boycott-assisting technologies more generally, we conducted an in-the-wild user study. To do so, we released Out of Site into the Chrome and Firefox Web Stores for participants to download. We then recruited people interested in our two proof-of-concept boycott campaigns to install Out of Site. We collected survey data, interview data, and log data from participants. The study was developed in concert with the research team’s local IRB and was eventually determined to be exempt due to relatively strict anonymization procedures and restraint in the data that was captured. Below, we first provide more detail about our two proof-of-concept campaigns and why they were selected. We next discuss the context and approach of our user study, which emphasized in-the-wild, exploratory observation. Finally, we present our quantitative and qualitative results, which are organized together into four themes that were established through an affinity diagramming process.

Proof-of-Concept Campaigns

For our user study, we implemented Out of Site campaigns for two existing boycott communities: GrabYourWallet and Stop Animal Testing. The GrabYourWallet boycott is a grassroots effort aimed at companies that have any connection or businesses with U.S. President Donald Trump and/or his family. The boycott has a list of targeted companies on its website and a team of organizers monitors the companies on this list and updates the list when necessary. The Stop Animal Testing campaign is led by People for the Ethical Treatment of Animals (PETA). This campaign highlights a list of cosmetic and household products that it suggests people avoid due to the animal testing that was used to develop these products . We selected these two campaigns because they have attracted a large number of participants. This meant that we would have a large potential population of extension users. While choosing this combination of campaigns provided the benefit of integrating with two well-established boycotts, this combination also presented a few challenges. In particular, the GrabYourWallet campaign is boycotting Amazon as a company (as well as many of the companies that sell products on Amazon.com). As such, in our first deployment phase, for users who accepted our default settings, their Amazon links in Google Search were filtered out and some of their visits to Amazon were redirected to the GrabYourWallet campaign website. This likely reduced users' visits to Amazon.com, which are already much lower than those to Google, and prevented us from gathering as much data about interaction with Amazon.com and the corresponding intervention types as we expected. In the case of our first deployment, the limited data we did gather from Amazon.com came from users who turned off the GrabYourWallet campaign, set its strength to "Medium" or "Low", went to Amazon.com more than once in an hour, or whitelisted Amazon.com. The installation wizard mitigated these issues in the second deployment, although interestingly it did not lead to a major increase in data from Amazon.

User Study: An Exploratory and In-the-Wild Approach

Following the deployment of Out of Site to the Chrome and Firefox stores, we sent out recruitment materials to existing online communities and message boards that are related to the GrabYourWallet and Stop Animal Testing boycotts. For instance, we advertised in Facebook groups and sub-reddits that are dedicated to women's rights issues and animal welfare. We also recruited members of our local community interested in these topics. As noted above, this paper reports the results of two separate phases of deployment. In between the two phases, we did extensive development based on feedback and log data from participants in this first phase, as described in the Design section. We use the term "first version" and "second version" to distinguish between the versions of the extension used in each phase. Our first deployment phase had 54 installations, with 42 people using the extension more than one day. This first phase lasted three weeks and average usage time was 6.7 days (although this was substantially attenuated by users who signed up midway through the first phase, leaving less time in the phase for usage). The second phase lasted four weeks and included 26 users from the first phase who had continued using the extension (the in-between phase data was not considered), as well as 19 new users who were recruited through a new round of advertising (21 new installations; two used the extension less than one day). The average usage length in the second phase was 10.0 days. We restrict our log analyses to people who used the extension for more than one day, although we analyze the available data on less-than-one-day users as well to gain a better understanding of non-use. Immediately after installation, users were asked to fill out an optional survey about their prior experiences supporting boycotts and other civic campaigns. The 48 users who completed the survey (36 from the first phase; 12 from the second phase) reported that, as predicted by the political consumption literature, many of them (46) had already been incorporating company ethics into their consumer behaviors and have had experiences with boycotting a variety of organizations, including the NRA (39), Uber (33), and Wal-Mart (18). In the first phase, the extension's default settings enrolled users into both the GrabYourWallet and the Stop Animal Testing campaigns, with both campaigns' strength level set to "High". However, users were able to customize their enrollment and settings freely and could turn off one or both of the campaigns easily. In the second phase, users were walked through these key settings in an installation wizard. The log data we collected through the extension was limited to two types: (1) how participants interacted with the browser extension itself, e.g. turning it on and off, whitelisting targets, and changing strength levels and (2) statistics from web pages that are affected by the extension. These log data were then uploaded to our database every 24 hours. In an effort to protect our users' privacy and in accordance with our IRB, our extension only collected information about visits to pages it had modified. Specifically, only altered Google SERPs, altered Amazon pages, and direct visits to targeted websites were recorded by the extension. For similar reasons, we also did not collect identifying information in the log; no experimenter is able to directly tie any specific user to their log data (although research has shown log data can be used for deanonymization with some effort [35]). To collect qualitative feedback, we reached out to participants via email with a request to interview them during the first deployment (e-mail addresses were not tied to log data). Seven users responded to our emails and were interviewed over phone or via text messages for approximately 30 minutes on average. Interviews were open-ended to elicit both generic feedback about the boycott-assisting technologies concept and granular feedback about Out of Site's settings and features. Each interviewee was compensated with a $10 electronic gift card, which was sent out via email at the end of the interview. Participants were permitted delete the extension from their browsers whenever they desired. When a user uninstalled the extension, an optional exit survey was shown to elicit any final feedback that users might have.

To understand our results, we first combined our (1) interview data (which we transcribed), (2) our survey data, and (3) written observations from an exploratory log data analysis. We then conducted a standard affinity diagraming process. Affinity diagramming is a popular approach among HCI researchers and practitioners to identify themes in heterogeneous data (e.g. [7,8,31,67]). Specifically, to execute the affinity diagramming, two members in our team created codes stemming from our three sources of data. We conducted two sessions of diagraming and iteratively refined our themes as new data became available. The final output of the affinity diagramming process was a set of four themes that cut across our quantitative log data and our qualitative survey and interview results. Data from our second deployment were combined into these four themes later to support or contrast with our previous findings. The four specific themes are: (1) Out of Site had a meaningful impact on users' web experiences, (2) there was a tension between automating action and automating awareness, (3) participants had positive reactions to collective action features, and (4) there was some non-conformance to the campaigns' goals.

Out of Site had a meaningful impact on users' web experiences

Out of Site affected 660 and 480 web pages in the first and second deployment phases, respectively. Across both deployment phases, the vast majority of the web pages were changed because of the GrabYourWallet campaign (655/660, 480/490). This distribution is not unexpected: the Stop Animal Testing campaign was a much more targeted campaign, only affecting Amazon and only a limited number of target companies (a portion of this effect is also likely due to the interaction between the campaigns discussed above). The GrabYourWallet campaign had a truly substantial impact on participants' Google search experiences; search engine results pages (SERPs) were the venue for the vast majority of the campaign's interventions (539 of 655 pages / 440 of 480 pages). Across both deployment phases, the filter intervention removed a total of 165 advertising links, 207 "knowledge graphs elements" (e.g. information boxes on the right side of search results [52]), 27 links to Twitter.com and over a thousand (1,065) standard search links. In the second deployment, Out of Site additionally filtered out 37 links to third-party commercial websites (e.g. Facebook, Trivago, coupon websites, Yelp). GrabYourWallet participants who used the "Medium" setting also saw consequential changes to their Google SERPs, although they were of the call-to-action intervention type rather than filter. Across both deployment phases, 214 search result links (of all types) were marked with a call-to-action on a total of 164 SERPs. We also saw a few users experiment with the call-to-action links that provide users an e-mail to send to the targeted companies. We only saw four users total use the "Low" setting in the GrabYourWallet campaign. Content-wise, the majority of the Google links affected were those to e-commerce websites. Amazon was by far the most impacted website. However, Macy's, Wal-Mart, Bed Bath and Beyond, Papa John's and Chewy.com were also affected in decent numbers. Indeed, examining the results, it is clear that Out of Site had its greatest effect for transactional queries [70] (e.g. "New Balance 530", "Papa John's"). This is the type of query most associated with commercial purchases in Rose and Levinson's three-part search query schema [70], meaning that Out of Site is having the intended effect of intervening in potential commercial transactions. Correspondingly, the SERPs that had fewer affected links usually were the result of "navigational" or "informational" queries, the two other types of queries in Rose and Levinson's schema. For instance, the query "chromebook video showing green" only had an Amazon link to a Chromebook affected. Similarly, the query "6.5 us to cm" resulted in a SERP that only filtered out a link to 6pm.com. Outside of Google, in the first deployment phase for GrabYourWallet, 116 direct visits to companies' websites were either redirected to GrabYourWallet's website (redirect intervention type; "High" setting) or blocked (block intervention type; "Medium" setting). The second phase saw a total of 40 redirected or blocked pages. Although the Stop Animal Testing campaign did not generate a large amount of log data, we observed that 40 animal testing products were affected on Amazon.com across 13 search queries in the first deployment phase. Similarly, 40 products across 10 search queries were affected in the second phase. As an example, one query for "Skinfood fresh fruit lip" resulted in the removal of products from ChapStick and Maybelline, two brands that were targeted by the campaign. Importantly, with respect to Out of Site's supporting of "buycotts", we observed evidence from both deployments that Out of Site diverted clicks to competitors of targeted companies. For example, in our first deployment, a user searched for “america’s first civilization michael coe” and received a SERP that had Amazon.com links removed. This user then actively engaged with multiple alternative search result links, including to Barnes Noble and abebooks.com (an online book store). Similarly, in our second deployment, a user searched for “Mihelcic and Zimmerman, 2nd edition” and received a SERP with filtered-out Amazon.com links. This user then visited a variety of online book stores such as abebook.com and wiley.com. With respect to our interview data, participants expressed almost exclusively positive opinions of the extension’s impact on their browsing experience (although this result is of course subject to observer-expectancy effects, novelty effects, and sampling bias). One interviewee (P4) tested the extension immediately after installation and was “excited to see it worked so well”. Similarly, another interviewee (P6) remarked with excitement that Out of Site is a “passively active approach” that “allows people make a social impact without having to do anything”. Another interviewee said of the GrabYourWallet campaign “It’s difficult to constantly keep track of all the businesses you interact with that oppose your values... It looks like [your] list is (automatically) updating.”

Tension between automating action and automating awareness

We saw significant evidence of users altering the strength settings (i.e.,“High”, “Medium”, “Low”) of their campaigns. In the first phase, five participants switched the GrabYourWallet campaign to “Medium”, one did so for Stop Animal Testing, and five users experimented with different strength settings and returned to the default “High” setting. In the second phase, because of our installation wizard, a number of participants

(7/19) chose to use “Medium” or “Low” settings (“Medium” was far more popular). From our interviews, it is clear that some users appreciated that the “High” setting automated their boycotting actions, while others felt that this setting was too invasive and instead wanted the extension to automate their awareness, i.e. by flagging targets when they encountered them. This latter group found the “Medium” setting to be the most effective. P7, a member of the group who remained on the “High” setting spoke very positively of its ability to automate actions: “I signed up for PETA’s mail list and followed a couple of advocates on Twitter. I usually look for information online, like looking into news articles, just Google it. I really like the Stop Animal Testing campaign… It (the extension) has been something I am looking for. It filters out things automatically.” Two of our interviewees set the extension to “Medium”. One of these participants (P2) remarked, “I want to know what is blocked and when, so I don’t miss anything important… the information will still be there if I need it.” Seen in an HPD light, this user benefited from and appreciated the intervention we implemented to prevent the filter bubble use case. The other participant (P1) who switched the extension to the “Medium” setting did so to become more familiar with the extension before trusting it to take action on her/his behalf: “I am a tinkerer. I used [the] ‘Medium’ setting to get familiar with what the extension does. I need some time to go on [the] ‘High’ setting.” Overall, these data reveal a tension in how to implement Out of Site’s boycott-assisting automation: some users appear to want actions automated (e.g. content filtered out). Others simply want the extension to automate the awareness process, helping them to understand when content from targeted companies is surfaced, but allowing them to take their own actions. For now, this suggests that Out of Site's implementation of a user-configurable setting is advisable, but there is likely more that can be done to navigate this tension. We return to this point below in Discussion.

Collective action features caused excitement

As predicted by the social computing collective action literature, participants reacted quite positively to Out of Site's collective action features. For instance, one interviewee (P6) echoed the findings in this literature that increased visibility of group progress incentivizes individual participation: "I like that information showing how many people joined you and how many products were hidden. It shows people are making progress." Similarly, another interviewee (P1) pointed out that the visibility of campaign progress could also help the campaign achieve its action-oriented goals: "If such thing snowballs, it could make a bigger impact. The effects of economic boycotts aren't often immediately visible. This could make the impact visible to large corporations." Furthermore, interviewees envisioned how the extension could be integrated with existing organizations and online communities to coordinate collective action: "I hope there would be a way I can communicate with others, like I can directly contact the organization if I have any questions." (P6). Similarly, other interviewees expressed the need to leverage social media to disseminate calls-to-action. As described above, this feedback was in part what motivated us to integrate social media features into the second version of Out of Site.

Non-conformance was observed

Participants in our study at times attempted to evade the automated boycotting assistance provided by Out of Site to meet an immediate need. To do so, some users employed the customization features that were provided by the extension. For instance, in our first deployment phase, out of five users who chose the "Medium" settings for the GrabYourWallet campaign, three of them changed their setting to "Medium" after failure to retrieve needed results from Google search. One interviewee (P5) emphasized what he believed to be the necessity of having the strength features in the extension for this purpose: "I think the simplicity of this extension itself is a good tool, but sometimes 'Medium' would be better if there's only one company making a thing on amazon, or if you just need to get some papa johns for whatever reason, especially because they have a big discount for college students." In addition to lower strength levels, users also employed the whitelist feature as a workaround for accessing targets. 17 users leveraged this feature in total across both deployment phases, 16 of whom whitelisted Amazon (other whitelisted targets include Papa John's, chewy.com, Bed Bath and Beyond, Wal-Mart, US Bank, and Belk.com). We also saw evidence that some of the users who uninstalled the extension (see above) did so for reasons related to non-conformance. For instance, in one of our six exit survey responses, one user mentioned that the extension blocks shopping websites and causes inconvenience: "The few times I needed to shop online I couldn't use normal sites at least not for a while." We also wondered if users who engaged with the extension for less than one day dropped out for the same reason (14 users in total), but were unable to infer more information from their log data. The majority of these users only opened the settings of the extension once and did not otherwise engage with the extension.

Discussion

Our deployments provided early evidence that Out of Site's vision of boycott-assisting technology has significant potential. Below, we first detail a number of implications for the design of boycott-assisting technology that arise from our research. Next, we discuss ideas for advancing our approach to HPD. Finally, we close by highlighting several important limitations of our research.

Design Implications

Customization Capabilities are Important: Our qualitative data suggests that participants valued and benefited from the customization capabilities included in Out of Site. In particular, in our user study, we observed that users have very different preferences between automating awareness and automating actions, and customization allowed them to adapt Out of Site to their preferences. Future work could take one step further and provide users with even more choices. For example, a feature could be added that allows users to set dates and times of participation, e.g. Mondays through Fridays or 10:00-14:00 every day, which may reduce campaign drop-out rates. SERPs are an Effective Site of Action: Another major insight from our user study is that the adaptation of search engine results pages (SERPs) can be a very effective mechanism for boycott campaigns. In our two short deployments, Out of Site made well over 1,000 changes to our participants' SERPs. Boycott-assisting technologies – and likely other collective action campaigns – are likely well-served by focusing development effort on contesting the information delivered in SERPs. Automation and Activism: At a high level, our user study provides evidence of the promise that automation holds for boycott-assisting technologies, as well as for activist and civic technologies more generally. Although our results suggest that this automation should be paired with significant user customization capabilities, it is clear from our qualitative data that the automation in Out of Site enabled our participants to action their values at scale in a way that they found empowering. Out of Site is part of a family of automation-assisted activist and civic technologies (e.g. ResistBot [**Res**]), and our results suggest that this family should grow. Such automation strategy has been widely applied to other fields such as computational journalism to monitor events and produce content [**Lokot and Diakopoulos2016**]. In a similar vein, activists could benefit from automation technology that helps to monitor social and political issues and take simple actions (e.g. advocating on social media). Community Functionality is Desired: Although Out of Site currently does not have a social aspect, multiple interviewees requested the ability to directly communicate with campaign organizers and other participants. Existing platforms that mobilize collective efforts such as change.org and gofundme.com provide a space for campaign participants to share their personal stories and motivations, as well as for campaign organizers to provide periodic feedback. It could be useful to integrate similar functionality either into Out of Site or on an associated website. How to Scale Up?: As discussed above, our HPD process revealed a major tension between the capability of Out of Site to support basically any boycott by any community and our desire to ensure that Out of Site is not used by hate groups and related organizations. This resulted in us pausing development of the easy-to-use wizard that outputs the JSON object that defines each campaign. Moving forward, it would be ideal to have this wizard in place and hosted online, but to couple this with the sociotechnical development of the gatekeeping process described above. This would make Out of Site resemble

Apple's App Store: like Apple's SDKs, our wizard would provide the capability to easily build a powerful tool, but we would also have a rigorous submission and approval process before the tool is launched. Another factor to take in consideration while scaling up such boycott-assisting technology is how to deal with conflicts between campaigns, e.g. the issue related to Amazon in our deployments. One way to address these types of issues is for users to be able to rank their campaigns by priority. Out of Site could then use these ranks to help users resolve these conflicts. Automated Keyword/Domain Identification: As noted above, one of few non-trivial tasks associated with building a campaign for Out of Site is identifying the set of keywords and domains associated with targeted companies' subsidiaries and brands. GrabYourWallet's official extension addresses this challenge by blocking all the domain names that are listed on their campaign website. However, we noticed that as many of the targets are conglomerates, such as Amazon, their subsidiaries sometimes are not flagged as targets in the extension. In the process of developing Out of Site, we identified conglomerates' subsidiaries using data from Wikipedia and public records. This was not much of a burden even for GrabYourWallet, which has many targets, and many of these are conglomerates. However, we did find that the manual approach was somewhat error-prone: in our first deployment, we accidentally missed a few of Amazon's subsidiaries (e.g. IMBD and Goodreads). Although this was easily fixable as keyword and domain tracking is managed on the server side, this also highlighted for us the importance of developing an automated keyword/domain generation tool. Using such a tool, an organizer could simply add the company names of which the organizer is aware and the relevant keywords and domains for all subsidiaries and brands would be output automatically. This is likely a tractable problem given the increasing availability of semantic web data (e.g. Wikidata) that contains subsidiary and brand relationships. Indeed, this problem is currently on the development shortlist for Out of Site. Addressing Non-Conformance: Above, we saw examples of participants who found ways around Out of Site's boycott guardrails to engage in activity not encouraged by the targeted boycotts. Prior work has found boycotts are most likely to succeed when purchasing the targets' goods or services is a highly visible action [**Friedman2002**]. This visibility is much weaker in online settings, where other participants cannot see a person walk into a targeted store or walk out of a store with a targeted item. Fortunately, by mediating the online experience, Out of Site is well-situated to address this downside of online boycotts. This will have to be done with care, however. To shame individual users publicly would almost certainly result in supporting our blacklisted use related to undue harm, although in a new way (as can be predicted by the online harassment literature [**Basak et al.2016, Jhaver, Chan, and Bruckman2018**]). One more positive approach might be to show an anonymized, aggregated statistic in our group progress display that indicates how many non-conforming visits and purchases have occurred in total. Replacement Discovery: Our interviewees reported cases in which alternative sites or products to those targeted by a boycott were not available, which is consistent with prior studies in the boycott literature [**Friedman2002**]. If alternatives are not available, boycott participants do not have any options other than purchasing from the boycott's targets. Fortunately, some campaigns have started aggregating alternative options to recommend to their participants. For example, PETA has a large database consisting of "cruelty-free" brands that do not use any animal testing based on their research. Ethicalconsumer.org provides a list of companies' ethical rating that takes multiple factors into consideration (e.g. environment, social responsibility). Out of Site could use these databases to power a recommender system that could be integrated into Amazon and Google search results as a new intervention type.

1. Advancing Heuristic Preventative Design

While we have enumerated how we utilized HPD for our project above, it is useful to briefly consider how it might be operationalized for other projects to explore its generality. We believe that for nearly all the research domains mentioned in the FCA proposal [**Hecht et al.2018**], HPD could provide useful insight and likely mitigate some negative impacts. For instance, in the case of the generation of audio and video with neural networks, HPD would likely result in a blacklist that includes the use of the neural networks to make propaganda. This would then encourage the research team to find ways to build watermarking or related approaches into the core of their approach (rather than treating it as a separate problem). Similarly, a research project that advances brain-computer interfaces might generate a blacklist that includes unwanted read/write access to specific parts of the brain. They would then work to prevent that use case within the initial contribution. Finally, a research team building a tool that semi-automatically tracks food consumption and encourages healthy eating would likely want to blacklist a use in which people with eating disorders co-opt the tool to advance their disorder. If HPD were to spread in popularity, a clear and important next step would be the creation of an aggregate list of well-motivated potential problematic uses of various types of computing innovations. This list would be an analogue to the standard heuristics used in heuristic evaluation [**Nielsen and Molich1990, Nielsen1994b**] and would serve as a key input to project-specific blacklists. Such a list would additionally further reduce the burden on researchers and developers employing HPD, potentially adding to HPD's broad accessibility. However, one challenge here would be adequate summarization and navigability, as this list could grow unwieldy over time.

Limitations

Although (and perhaps because) Out of Site advances social computing's understanding of a new problem space – automation-assisted boycotts – our research is subject to some limitations. First, as is often the case with deployed systems, Out of Site had a few bugs, especially in the first deployment phase. For instance, as noted above, for the first-phase GrabYourWallet campaign, IMBD and Goodreads were not flagged as subsidiaries of Amazon (this was fixed for the second phase). It is also worth noting that some of our search queries on Google or Amazon might be users merely experimenting with the extension. We do not expect, however, that either of these issues had a meaningful impact on our exploratory user study and its high-level observations. Another limitation of the study is that we tracked a very small fraction of users' browser histories in an effort to protect users' privacy. As such, we were unable to identify additional websites or content that could have been targeted (e.g. it could be that supporting Wal-Mart's search function is important for, for instance, GrabYourWallet members). It is also important to note that as both proof-of-concept campaigns share somewhat similar low-level political ideologies, our user study's result might not apply to other demographics, as different ideological groups may adopt different tactics (e.g. [**Hond and De Bakker2007**]).

Conclusion

In this paper, we have described Out of Site, a boycott-assisting technology that automates many of the challenging aspects of implementing successful boycotts. We described the unique design approach we took with Out of Site that we call heuristic preventative design and reported on the use of Out of Site in two deployments with 42 users and 45 users, respectively. We observed that Out of Site substantially changed users' web experiences and that some users preferred to have their actions automated while others simply wanted assistance with awareness of relevant information. We also observed some attempts at non-conformance with respect to boycott goals. Our results support the strong potential of Out of Site and boycott-assisting technologies more generally and inform means by which boycott-assisting technologies can meet this potential.

Chapter VI: Prioritizing Data Producers

In the previous chapters, I have focused on understanding and characterizing the present landscape of data and data-driven technologies. But to enact systematic changes to the data economy, guiding principles are necessary to orient the collective efforts of researchers, practitioners, policymakers, and activists. In particular, Chapter Two lays out the immediate, short-term goals that revolve around empowering data producers in different contexts (e.g. visible vs. invisible data labor), there is still a lack of systematic blueprints for the data economy as a whole. A long-term vision for the data economy would help to inform action plans and priorities for all stakeholders so that the future of the data economy is built upon shared values, goals, and public interests.

My last chapter will draw from related literature and my earlier work to envision principles for an alternative data future that prioritize data producers instead of companies. In this chapter, I asked, who are the stakeholders of the data economy? What if we envisioned an alternative data future that prioritized data producers' interests over those of other roles? What shared principles and values might this future uphold? My ultimate goal is to pave the way for researchers, activists, and policymakers to build data futures center data producers' interests. I drew from extant research that focuses on governing and monetizing data from two areas–data governance and data economics and mapped out the roles and principles supported by these two bodies of work. Extending my prior taxonomy of data labor, I also reviewed literature on worker-centered design to inform additional principles that might apply to the data economy. I focused on addressing two specific research questions:

RQ1: What existing roles are considered as the stakeholders of the data economy?

RQ2: What principles might the future data economy be built upon?

My results show that the data governance and data economics literature primarily focus on data-driven businesses' interests and proposing guidelines that can maximize data's benefits for businesses while minimizing privacy risks to data producers. Only recently, scholars have expanded to advocating for guidelines that are in the interest of data producers beyond privacy, such as fair compensation and control. My review of worker-centered design literature provides additional guidelines that highlight opportunities to further engage with those that provide data labor in the design and development of data-driven technologies. Together my literature reviews synthesize principles that data producers may pursue in the future so that data-driven technologies are reflective of data producers' values and interests, not just those of businesses. I will further discuss how researchers may accelerate this process, by providing research-based guidance and designing new data governance frameworks that enforce data producer-centered principles.

In the rest of the chapter, I will first outline how I conducted my scoping reviews of literature from the three areas—data governance, data economics, and worker-centered design. I will then report my findings from the literature review to answer RQ1 and RQ2, respectively. Lastly, I will discuss promising research opportunities to actualize these data producer-centered principles.

Methods

The three research areas I have chosen above have widely varied amounts of literature. Data governance is an established field of research with a wide range of sub-areas such as cloud computing and database management. Conversely, data economics and worker-centered design are two growing areas of research and therefore have limited amount of literature. Given the differences in the quantity of literature among these three areas, I used two different survey methods to conduct my literature reviews. Specifically, for the data governance literature, because of the vast amount of research already accomplished in the area, I primarily synthesized literature reviews on data governance. For the other two research areas–data economics and worker-centered design, I conducted a scoping review of literature by searching for related terms as detailed below.

Data Governance - A survey of literature reviews

Data governance has historically attracted attention from a variety of scholars, industry practitioners, and policymakers. As a result, data governance principles represent the interests and perspectives from a diversity of roles and groups, including businesses, policymakers, and researchers.

There exists a large amount of literature on data governance and multiple systematic literature reviews on the topic, with slightly varied focuses. I focused on synthesizing systematic literature reviews to understand what principles are recommended in the context of data governance. I used keywords "data governance" and "literature review" to search for related papers on Scopus—a large database of scientific articles. I then filtered out articles that are not closely related to data governance, yielding a total of 20 articles. To answer RQ1, I first identified what stakeholders are referenced in these articles. I then extracted principles and values that these articles recommend to shed light on RQ2.

Data Economics

As data transforms business models and leads to technological advancements, the financial characteristics of data have become an important area of research for economists, computer scientists, and practitioners. I conducted a scoping review of literature on data economics, focusing on two important sub-areas: data valuation and data sharing. As the research area is fast-evolving and growing, I experimented with a variety of keywords to retrieve relevant studies, including "data pricing", "economics of data", "data price", and "data market". In the end, I classified 32 papers that provide the most up-to-date understanding about data economics.

In answering RQ1, I identified who are the stakeholders involved in data pricing. In answering RQ2, I focused on what guiding principles are used to build data models. Or put another way, what values did scholars instill in their data valuation and data sharing models?

Worker-Centered Design

As computing systems become widely adopted across industries and sectors, worker-centered design has been advocated by HCI researchers to empower workers in their day-to-day employment. As early as 1996, Greenbaum argued for a shift of focus to labor process in studying and design social computing systems. Scholars' interests in labor have persisted over the years, and expanded from traditional types of labor to emerging forms of labor such as digital labor and creative labor.

My survey of literature in this space largely relied on papers from prominent HCI publication venues. My goal was to identify potential avenues of support for those that provide the fundamental data labor for the production of technologies. Or put another way, what principles of worker-centered design might be applicable to data labor? As worker-centered design is an emerging area of research, scholars have not yet coalesced on a set of terms, posing a challenge for my survey. I extracted a subset of papers from the pool of labor-oriented papers mentioned in Chapter Two that focused on examples of data labor (e.g. producing user-generated content) and advocated for worker-centered design approaches and 22 out of 78 papers met this criterion. I supplemented this set of papers with prominent studies cited by prominent worker-centered design essays and proposals, such as Fox et al. [**Fox et al.2020, Fox, Sobel, and Rosner2019**].

Results

RQ1: What roles are considered the stakeholders of the data economy?

The data governance literature and the data economics literature have identified and investigated a set of different roles that have a stake in the topic. Below, I provide the prominent descriptions of these roles, in answering RQ1.

Organizations that collective and/or house data is a major focus for the body of work on data governance. Authors have proposed different frameworks for organizations to adopt with respect to facilitating their day-to-day operations around data governance (e.g. [**Al-Ruithe, Benkhelifa, and Hameed2019**]). These organizations may vary, from higher-education institutions to IT businesses. Work on this subject often concerns how to govern data within organizations and between organizations and assumes organizations as the entities with some if not all decision making power around data, such as who has access and how to share data.

A portion of the data governance work proposed and constructed the role of data stewards and data trustees [**Janssen et al.2020, Carroll et al.2020, Griffiths et al.2021**], who provide oversight over how data is used. Data stewards are proposed in part to address the ambiguity of data ownership. Specifically, because data can be shared freely with little cost, data ownership is often challenging to discern, if at all possible. The concept of data stewards emerged in response to this call as a shared, collective approach to data governance instead of a individualistic approach. However, while widely proposed in literature, data stewards lack clear-cut examples in reality. While some data in the public domain may be managed and supervised by certain public institutions, such as open data office in municipal governments and science data archives, these entities have limited capabilities in overseeing the use of the data under their stewardship. Exploring and strengthening data stewards' role in the data economy is, therefore, an important next step to redistribute the decision making power around data from companies to the public.

Additionally, perhaps unsurprisingly, researchers are a prominent type of stakeholders in the data governance literature. Much of the work is in conversation with researchers to identify new research opportunities, close research gaps, and innovate new frameworks. And similarly, practitioners and developers are identified as a key stakeholder in data governance as they are the ones that implement technical infrastructures to facilitate the production, aggregation, and utilization of data.

In data economics, there are two distinct roles involved in the data marketplace: data sellers and data buyers [**Bergemann, Bonatti, and Gan2021**]. The distinction between data sellers and data producers is at times unclear. Data sellers could refer to those that host data in their servers but do not produce data. For example, much of the work on data pricing for ML and cloud computing refers to the host of data as data sellers [**Pei et al.2021**]. But data sellers may also refer to those who produce data and are looking to trade the corresponding data on a marketplace, as Acemoglu et al. simulated [**Acemoglu et al.2019**].

The data economics literature also sheds light into some examples of data buyers. Much of work in this space emerged from pricing big data for ML and AI, so companies that develop and sell ML services are a key type of data sellers [**Cong et al.2022**].

RQ2: What principles might the future data economy be built upon?

The three research areas I surveyed have drastically different perspectives with respect to the principles of the data economy. While data governance scholars sought to be all-encompassing to support values and interests of a diversity of stakeholders, work on data economics largely focus on optimizing the data marketplace to maximize the benefits of data for both data buyers and data sellers. Worker-centered design has taken on the perspective of those that provide the labor for computing systems. Below, I list the principles that emerge from all three areas.

Privacy: privacy is a concern that cross-cuts data governance and data economics. For example, Choi, Bergemann, and Acemoglu have all independently sought to minimize privacy in their work on simulating data market mechanisms [**Choi, Jeon, and Kim2019, Bergemann, Bonatti, and Gan2021, Acemoglu et al.2019**]. Privacy is also a key concern in the data governance literature. The majority of the literature reviews explicitly listed privacy as a core goal for those that manage data including organizations and practitioners.

Transparency/information symmetry: Jassen et al. argued that transparency in data governance is crucial to enable external oversight and critique so that companies will responsibly and ethically use data [**Janssen et al.2020**]. More over, Acemoglu et al. conducted simulations of data markets and shown that transparency of data collection and use would also assist data producers in making informed decisions about the implications of data at the time of production [**Acemoglu et al.2019**]. Providing data producers with the transparency about how their digital traces are used downstream by other actors is a key area of improvement for the current data infrastructure.

Control and agency: Jassen et al. listed control of data as a key principle to implement in data governance [**Janssen et al.2020**]. Similarly, Abraham took the perspective of organizations and prompted future work to investigate "how do organizations retain control over their data?" [**Abraham, Schneider, and vom Brocke2019**]. These two examples highlight two approaches to exerting control and agency: individual control and collective control. Privacy and security researchers and practitioners have made great strides in understanding how to establish meaningful, individual control over one's data. Conversely, discussions and ideations about collective control remain conceptual, such as data cooperatives and data trust. This is one promising area for researchers and data producers to collaboratively experiment with, so that aggregated data is used in accordance with shared, collective values and goals.

Collective benefits: Equitable benefits for all data producers is a goal that has been highlighted in the governance of indigenous data[**Carroll and Bellotti2015**]. In the domain of health research, Griffiths et al. has proposed specific ways that may ensure data has collective benefits for indigenous communities, such as hosting their own health information systems and setting up councils to oversee the use of indigenous data [**Griffith2017b**]. This goal is not only applicable to indigenous data but also has been discussed by scholars and policymakers across the globe. For example, various policy proposals on data trusts and data dividends have emerged across the globe, e.g. the European Union's data governance act and California's data dividends (Lohr).

Responsibility and accountability: Janssen et al. and Carroll et al. both advocated for having formal responsibility and accountability mechanisms that mediate how data is used [**Jhaver, Bruckman, and Gilbert2019b, Carroll and Bellotti2015**]. This could mean having a special role such as data stewards to oversee data collection, aggregation and utilization. Moreover, Felici et al. argued that accountability in data systems is crucial to gain trust from data producers [**Felici, Koulouris, and Pearson2013**]. As the landscape of data governance rapidly changes, accountability mechanism is a key area that researchers, practitioners, and policymakers should all focus on to create incentives for responsible data use and deter data practices that lead to negative impacts on data producers.

Arbitrage-free: Arbitrage-free pricing means that data transitions have a consistent value across different contexts and market channels [**Pei2020**]. Or put another way, data buyers will have to purchase data at a uniform price. While this principle is not always be guaranteed in real life contexts, it is ideal to minimize pricing discrepancies so all data sellers will be compensated fairly.

"Good" jobs for data labor: This means an environment or online space in which data producers can contribute high quality data with minimal burden or risk. Currently, some types of data labor are labor-intensive or even harmful for data producers, e.g. content moderation work and crowdwork [**Gray and Suri2019**]. Improving the working condition of data producers may mean improved user experience, meaningful compensation for labor hours, and customized supports for those that produce specialized data. In particular, those that produce the most labor should be prioritized and equipped with adequate support for their work. Using advertising as an example, the attention of those that produce data is a good that technology companies are selling to ad buyers. We should shift this way of thinking to treating data producers as workers. Subsequently, resources and supports such as better tools and working conditions should be provided to workers.

Work(er) visibility: Studies of work, workplaces, and modern digital labor have consistently shown how digital traces shift the boundaries between visible and invisible work and subsequently affect workers' status (e.g. [**Star and Strauss1999, Suchman1995a**]). Making data labor(ers) visible therefore highlights the collective nature of aggregated data and acknowledge the important role data producers play in powering some of the most powerful data-driven technologies.

Fair compensation: Multiple studies from data economics have advocated fair compensation for data in order to establish a market that rewards quality and cultivates innovation [**Cong et al.2022, Pei et al.2021, Agarwal, Dahleh, and Sarkar2019**].

In particularly, Pei's review of data pricing highlighted Shapley value may be a promising approach to value data so payment can be fairly distributed among all data sellers [**Pei2020**]. It is worth noting, however, that these studies operate from the perspective of data sellers, i.e. groups that have aggregated datasets, rather than individual data producers. Future work may extend extant work and further investigate how data sellers may divide data compensations to individual payments. Moreover, issues around compensation for labor such as wage theft and low wages is also a focal point of research for worker-centered design and researchers have highlighted the burden workers take on in their day-to-day practices, ranging from emotional labor to lost wages. While the vast majority of data labor is uncompensated, there may be other forms of compensation that data producers may pursue, such as protection from online harms, supports for emotional wellbeing, and reimbursement for products and services used for data production.

Discussion and Future Work

Involving data producers in decision making about data

While literature on data governance and economics accounted for a diversity of roles in the data economy, data producers were rarely included. As a result, much of the prior frameworks of data governance and marketplaces failed to take in data producers' perspectives and incentives. As data is growing exponentially, how to involve data producers in decision making about data is one urgent area of research. Future work may answer questions such as how might researchers and practitioners seek consent from data producers when building data-driven technologies? How might data producers limit the downstream use cases of data as they wish? And what accountability mechanisms could be helpful to preserve data producers' rights to data?

While a step towards the right direction, involving data producers in decision making about data will likely face practical challenges. One key challenge is data producers' potential conflicts with other stakeholders of the data economy with respect to goals and interests. For example, while data-driven platforms such as Facebook and Twitter seek to grow their user base, this goal does not necessarily align with their existing data-producing users. To mitigate this challenge, the literature on value-sensitive algorithm design and participatory AI may be particularly helpful. Prior work in these areas have shed insights into how to navigate and address the conflicts in goals and preferences among a diversity of stakeholders. Future work on data governance may apply similar approaches to include data producers as a key stakeholder in the design and development of data-driven technologies.

What principles are of high priority for data producers

Future work may pick up where I left with these principles and investigate which ones are of high priority to data producers. Given that data production can vary widely depending on the social and economic contexts, what may be of high importance for a group of data producers may not be so for other groups. More specifically, in crowdsourced data production, worker well being and fair wage are likely to be what data producers i.e. crowdworkers value the most. Whereas in more sensitive contexts such as producing data about mental health disclosure, privacy is likely to be the top concern of data producers. Because of the diverse ways in which people contribute data, future work needs to account for the varied incentive structures and contexts and test how different populations may rank these principles differently.

Moreover, future work may further investigate how different collective decision-making mechanisms such as sparse sampling, voting, and delegation may play out among data producers with respect to data use. Such work would yield empirical evidence about the strengths and weaknesses of different mechanisms for decision-making about data use. This work would also provide concrete guidance on transforming equitable data futures into tangible experimentations and practices. Currently, discussions and ideations of data governance primarily operate at a conceptual level. Such work will provide immediate guidance for developing data governance models that are driven by collective values and shared goals.

Designing for data transparency

One low-hanging fruit among these principles is transparency. A plethora of research, policymaking and advocacy efforts have laid the ground work for establishing data transparency, from GDPR to affirmative consent to data feminism. Future work may leverage such prior work and explore how to effectively and meaningfully make data's various use cases transparent to its producers. In particular, as current approaches to informed consent and terms of service have been largely proven to be cumbersome or senseless to data producers, how might we create new forms of data transparency? One area of literature that may be particularly informative is algorithmic fairness - researchers have developed various effective communicative techniques to illustrate the inner working of complex algorithmic models. Such techniques may be applicable to data use so that data producers can directly see the incremental impact of data points on data-driven technologies.

Future work may also take one step further and investigate data transparency's implications on data producers' attitude and behavior. This goal might be achieved by deploying tools that are similar to Out of Site and directly communicate the value and use cases of data points generated by data producers in real time. Researchers could then recruit participants to solicit their reflections on questions around how their data-producing behaviors might be affected by such transparency and what incentive mechanisms might be suitable for encouraging certain types of data production.

Re-imagining alternative models of data valuation

More broadly, future work may take these principles further by challenging the propertarian framing of data as property and seek to construct alternative models of data valuation that center these principles in lieu of monetization. Every day, billions of people generate troves of data with the technologies we use the most, from search engines to social media platforms and these traces of human activity represent our collective, digital, data cultures. But most data is collected, aggregated, and stored by just a small set of large, for-profit companies such as Google and Facebook. In a society where data is treated as private property, those who collect and keep data define its value and benefit from it financially. But what if we envisioned alternative models of valuation that featured the social and collective labor of the people who create data instead of the platforms that collect and profit from it?

Re-imagining data valuation models will set the stage for alternative data futures that are driven by collectively defined values rather than solely monetization. Scholars, policymakers, and practitioners have extensively debated the future of the data economy; much of these debates and ideations are centered around monetary valuations of data. However, data's potential is much more than generating capital. Future work may expand Chapter 3's method and further recognizes all types of data production as a form of collective, social labor that produces data as a public good. In doing so, such work will highlight how online participation is largely enabled and subsidized by public institutions. Currently, discussions of data labor largely take on the propertarian framing in that data labor leads to individual assets with designated owners. But what if data labor assumed communities of creators and stakeholders? More specifically, researchers may interview data producers and data stewards-alike, including librarians, open data offices in municipal governments, and community data organizations to understand the hidden cost in the creation, collection, aggregation, and preservation of data.

Innovating compensation mechanisms for data labor

As companies innovate new ways to monetize data labor and accumulate wealth, it is equally important for policymakers and researchers to innovate compensation mechanisms for data labor. How companies should reward data laborers for their contribution to their business model is a question that has generated extensive debates among researchers across disciplines (Kugler). While one may argue that currently many data-driven companies are already rewarding their data laborers by providing free or low-cost services, it is unclear whether data laborers are receiving the short end of the stick. Moreover, my prior work has suggested that the amount of data labor subsidizing for-profit businesses plays a non-trivial role in upholding a company financially. Compared with traditional industry sectors such as retail, the labor share of the data economy is substantially lower, due to the fact that much of the data labor remains unaccounted for. Taken all together, it is crucial that we start to innovate compensation mechanisms for data labor so that those that produce fundamental materials, i.e. data, for data-intensive businesses receive support and incentives accordingly.

Broadening data access

The amount of data that we collectively generate has been increasing rapidly in the past decade. It is forecasted that by 2025, the amount of data the whole world produces will double the yearly estimate for 2021, reaching 180 zettabytes. Keep in mind that a zettabyte equals one billion terabytes. So if we put all these data in those 1TB hard drives that we use for backups and divide them among the world population, everyone on earth would get more than 20 hard drives.

Although one might have quite a few hard drives in their drawer, at a global level, the vast majority of data is kept on private servers and inaccessible to the public (sometimes for good reasons such as privacy protection). Moreover, a significant share of data is stored and accessed by just a small set of large, for-profit companies such as Google, Amazon, Facebook, and Microsoft. In a sector where data is treated as private property of those that collect them, technology companies are equipped with unchecked power to make data-driven decisions, deploy predictive models, and shape public lives. For example, one's profile picture may be used without their knowledge to train facial recognition technologies that put others under privacy and surveillance risks. Similarly, language models trained with data encoded with human biases may exacerbate such existing biases on a larger scale. The lack of easily accessible open data is also a the root cause for raising the bar-to-entry for data-intensive startup businesses and stifling technology innovation. Taken all together, increasing access to data, a collectively created, valuable resource, is the first step towards building a transparent, just data future for all.

Conclusion

By synthesizing related literature from three domains–data governance, data economics, and worker-centered design, I highlighted the lack of involvement of data producers in the decision making of technology development. I further identified a set of principles that a data producer-centered technology ecosystem should be built upon, charting out potential paths for the future of data labor.

Conclusion

In this dissertation, using surveys, data analysis, system testing, and literature reviews, I have shown how reconceptualizing data production as a form of labor can pave the way for data producers, researchers, activists, and policymakers to mitigate the power inequity between the public and data-driven technology companies. Based on my findings, I have proposed concrete recommendations for all stakeholders of technology to actively shape the design and development of data-driven technologies.

In particular, this dissertation highlighted two promising directions to foster a more equitable relationship between technology companies and data producers: raising awareness of the value of data labor and amplifying the collective voice of data laborers.

Understanding the value of data labor is the first step toward ensuring data-driven technologies have broad, collective benefits. As such, mapping out how various types of data labor are supporting tech businesses will inform concrete policy recommendations and data practices for policymakers, researchers, and practitioners. Chapter Four provided one approach to measure the value of data labor and future research and policymaking efforts may extend this approach to other types of data labor more broadly.

Organizing collective action among data producers is another direction that may lead to meaningful progress in changing current data practices and data-driven technologies. As shown in Chapter One, data producers are already engaging with protests against data-driven technologies individually. Aggregating their impact through systems like Out of Site has the potential to unlock data producers' collective labor power and subsequently affect tech companies' data practices.

Lastly, in studying and characterizing data labor, I have explored the dimension of visibility extensively in Chapter Three and Four; the remaining four dimensions remain under-explored. One particularly urgent area of research for future work is to improve end-use awareness. As new technological innovations like AI occur, data labor's downstream impact is becoming more and more nebulous to the public, raising the urgency for researchers to design measures that promote transparent and democratic data use. In my future work, I plan to further explore the end-use awareness dimension by making transparent the downstream use cases of data and designing collective decision-making processes for data producers and other data stakeholders.

References

[aaa2020] 2020. Towards responsible ai in surveillance, media, and security through licensing.

[Abraham, Schneider, and vom Brocke2019] Abraham, R.; Schneider, J.; and vom Brocke, J. 2019. Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management* 49:424–438.

[Acemoglu and Autor2011] Acemoglu, D., and Autor, D. 2011. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics*, volume 4. Elsevier. 1043–1171.

[Acemoglu et al.2019] Acemoglu, D.; Makhdoumi, A.; Malekian, A.; and Ozdaglar, A. 2019. Too Much Data: Prices and Inefficiencies in Data Markets.

[ACM] ACM Ethics. https://ethics.acm.org/.

[Agarwal, Dahleh, and Sarkar2019] Agarwal, A.; Dahleh, M.; and Sarkar, T. 2019. A marketplace for data: An algorithmic solution. In *ACM EC 2019 - Proceedings of the 2019 ACM Conference on Economics and Computation*, 701–726. Association for Computing Machinery, Inc.

[AIG2018] 2018. AI at Google: Our principles. https://www.blog.google/technology/ai/ai-principles/.

[Al-Ruithe, Benkhelifa, and Hameed2019] Al-Ruithe, M.; Benkhelifa, E.; and Hameed, K. 2019. A systematic literature review of data governance and cloud data governance. *Personal and Ubiquitous Computing* 23(5):839–859.

[Albrechtslund2008] Albrechtslund, A. 2008. Online social networking as participatory surveillance. *First Monday* 13(3).

[Anderson, Sarkar, and Palen2019] Anderson, J.; Sarkar, D.; and Palen, L. 2019. Corporate editors in the evolving landscape of openstreetmap. *ISPRS International Journal of Geo-Information* 8(5):232.

[Andorfer and Liebe2012] Andorfer, V. A., and Liebe, U. 2012. Research on Fair Trade Consumption—A Review. *Journal of Business Ethics* 106(4):415–435.

[App] App Store Review Guidelines - Apple Developer. https://developer.apple.com/app-store/review/guidelines/#objectionable-content.

[Arendt1958] Arendt, H. 1958. The human condition.

[Arrieta-Ibarra et al.2018] Arrieta-Ibarra, I.; Goff, L.; Jiménez-Hernández, D.; Lanier, J.; and Weyl, E. G. 2018. Should we treat data as labor? moving beyond" free". In *aea Papers and Proceedings*, volume 108, 38–42.

[Atkinson2012] Atkinson, L. 2012. Buying In to Social Change: How Private Consumption Choices Engender Concern for the Collective. *The ANNALS of the American Academy of Political and Social Science* 644(1):191–206.

[Au-Yeung2019] Au-Yeung, A. 2019. California Wants To Copy Alaska And Pay People A 'Data Dividend.' Is It Realistic? https://www.forbes.com/sites/angelauyeung/2019/02/14/california-wants-to-copy-alaska-and-pay-people-a-data-dividend–is-it-realistic/.

[Bandy and Vincent2021a] Bandy, J., and Vincent, N. 2021a. Addressing" documentation debt" in machine learning research: A retrospective datasheet for bookcorpus. *arXiv preprint arXiv:2105.05241*.

[Bandy and Vincent2021b] Bandy, J., and Vincent, N. 2021b. Addressing "Documentation Debt" in Machine Learning Research: A Retrospective Datasheet for BookCorpus. *35th Conference on Neural Information Processing Systems*.

[Barnett et al.] Barnett, C.; Cloke, P.; Clarke, N.; and Malpass, A. Consuming Ethics: Articulating the Subjects and Spaces of Ethical Consumption. *Antipode* 37(1):23–45.

[Basak et al.2016] Basak, R.; Ganguly, N.; Sural, S.; and Ghosh, S. K. 2016. Look Before You Shame: A Study on Shaming Activities on Twitter. In *Proceedings of the 25th International Conference Companion on World Wide Web*, WWW '16 Companion, 11–12. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

[Bates et al.2019] Bates, O.; New, K.; Mitchell-Finnigan, S.; Mauriello, M. L.; Remy, C.; Bendor, R.; Mann, S.; Chopra, S.; Clear, A. K.; and Preist, C. 2019. Towards a responsible innovation agenda for hci. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–8.

[Baumer et al.2013] Baumer, E. P.; Adams, P.; Khovanskaya, V. D.; Liao, T. C.; Smith, M. E.; Schwanda Sosik, V.; and Williams, K. 2013. Limiting, Leaving, and (Re)Lapsing: An Exploration of Facebook Non-use Practices and Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 3257–3266. New York, NY, USA: ACM.

[Baumer et al.2015a] Baumer, E. P. S.; Burrell, J.; Ames, M. G.; Brubaker, J. R.; and Dourish, P. 2015a. On the Importance and Implications of Studying Technology Non-use. *interactions* 22(2):52–56.

[Baumer et al.2015b] Baumer, E. P.; Ames, M. G.; Burrell, J.; Brubaker, J. R.; and Dourish, P. 2015b. Why study technology non-use? *First Monday* 20(11).

[Baumer et al.2015c] Baumer, E. P.; Guha, S.; Quan, E.; Mimno, D.; and Gay, G. K. 2015c. Missing photos, suffering withdrawal, or finding freedom? How experiences of social media non-use influence the likelihood of reversion. *Social Media+ Society* 1(2):2056305115614851.

[Baumer2015] Baumer, E. P. 2015. Usees. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, 3295–3298. New York, NY, USA: ACM.

[Baumer2018a] Baumer, E. P. S. 2018a. Socioeconomic Inequalities in the Non Use of Facebook. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, 616:1–616:14. New York, NY, USA: ACM.

[Baumer2018b] Baumer, E. P. 2018b. Socioeconomic inequalities in the non use of facebook. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–14.

[Baumgartner et al.2020a] Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020a. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, 830–839.

[Baumgartner et al.2020b] Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020b. The Pushshift Reddit Dataset. *Proceedings of the International AAAI Conference on Web and Social Media* 14:830–839.

[Becker and Copeland2016] Becker, A. B., and Copeland, L. 2016. Networked publics: How connective social media use facilitates political consumerism among LGBT Americans. *Journal of Information Technology & Politics* 13(1):22–36.

[Beecher et al.2019] Beecher, S.; van der Drift, D.; Martin, D.; Vass, L.; Goder, S.; Lim, B.; and Langner, M. 2019. Facebookincubator/bootstrapped. https://github.com/facebookincubator/bootstrapped.

[Bennett and Segerberg2012] Bennett, W. L., and Segerberg, A. 2012. The Logic of Connective Action. *Information, Communication & Society* 15(5):739–768.

[Bergemann, Bonatti, and Gan2021] Bergemann, D.; Bonatti, A.; and Gan, T. 2021. The Economics of Social Data. *arXiv:2004.03107 [cs, econ, q-fin]*. arXiv: 2004.03107.

[Bevensee et al.2020] Bevensee, E.; Aliapoulios, M.; Dougherty, Q.; Baumgartner, J.; McCoy, D.; and Blackburn, J. 2020. *SMAT: The Social Media Analysis Toolkit*. US: ICWSM.

[Bigham2018] Bigham, R. 018. It's easy to think that computer science will only have positive broader impacts, or that how it's used isn't your problem: Neither is true. my @acm_fca colleagues and I think we should treat Negative Impacts seriously when writing papers & seeking fundinghttps://acm-fca.org/2018/03/29/negativeimpacts/ ....

[Bødker2006] Bødker, S. 2006. When second wave hci meets third wave challenges. In *Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles*, 1–8.

[Boulstridge and Carrigan2000] Boulstridge, E., and Carrigan, M. 2000. Do consumers really care about corporate responsibility? Highlighting the attitude—behaviour gap. *Journal of Communication Management* 4(4):355–368.

[Bozarth and Budak2017] Bozarth, L., and Budak, C. 2017. Social Movement Organizations in Online Movements. SSRN Scholarly Paper ID 3068546, Social Science Research Network, Rochester, NY.

[Brown et al.2020a] Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; Agarwal, S.; Herbert-Voss, A.; Krueger, G.; Henighan, T.; Child, R.; Ramesh, A.; Ziegler, D. M.; Wu, J.; Winter, C.; Hesse, C.; Chen, M.; Sigler, E.; Litwin, M.; Gray, S.; Chess, B.; Clark, J.; Berner, C.; McCandlish, S.; Radford, A.; Sutskever, I.; and Amodei, D. 2020a. Language Models are Few-Shot Learners. *arXiv:2005.14165 [cs]*.

[Brown et al.2020b] Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. 2020b. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

[Bureaua] Bureau, U. C. 2017 National Population Projections Tables. https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html.

[Bureaub] Bureau, U. C. Income and Poverty in the United States: 2017. https://www.census.gov/data/tables/2018/demo/income-poverty/p60-263.html.

[Buy] Buycott. https://www.buycott.com/.

[Carrigan and Attalla2001] Carrigan, M., and Attalla, A. 2001. The myth of the ethical consumer – do ethics matter in purchase behaviour? *Journal of Consumer Marketing* 18(7):560–578.

[Carroll and Bellotti2015] Carroll, J. M., and Bellotti, V. 2015. Creating Value Together: The Emerging Design Space of Peer-to-Peer Currency and Exchange. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, 1500–1510. New York, NY, USA: Association for Computing Machinery. event-place: Vancouver, BC, Canada.

[Carroll et al.2020] Carroll, S. R.; Garba, I.; Figueroa-Rodríguez, O. L.; Holbrook, J.; Lovett, R.; Materechera, S.; Parsons, M.; Raseroka, K.; Rodriguez-Lonebear, D.; Rowe, R.; Sara, R.; Walker, J. D.; Anderson, J.; and Hudson, M. 2020. The CARE Principles for Indigenous Data Governance.

[ccp] California consumer privacy act (ccpa).

[Center2019] Center, P. R. 2019. Public knowledge and experiences with data-driven ads.

[Centivany and Glushko2016] Centivany, A., and Glushko, B. 2016. "Popcorn Tastes Good": Participatory Policymaking and Reddit's. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 1126–1137. New York, NY, USA: Association for Computing Machinery.

[Chancellor, Baumer, and De Choudhury2019] Chancellor, S.; Baumer, E. P.; and De Choudhury, M. 2019. Who is the" human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):1–32.

[Chancellor et al.2016a] Chancellor, S.; Lin, Z.; Goodman, E. L.; Zerwas, S.; and De Choudhury, M. 2016a. Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, 1171–1184. New York, NY, USA: Association for Computing Machinery.

[Chancellor et al.2016b] Chancellor, S.; Pater, J. A.; Clear, T.; Gilbert, E.; and De Choudhury, M. 2016b. #thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, 1201–1213. New York, NY, USA: Association for Computing Machinery.

[Chancellor et al.2019a] Chancellor, S.; Birnbaum, M. L.; Caine, E. D.; Silenzio, V. M.; and De Choudhury, M. 2019a. A taxonomy of ethical tensions in inferring mental health states from social media. In *Proceedings of the conference on fairness, accountability, and transparency*, 79–88.

[Chancellor et al.2019b] Chancellor, S.; Nitzburg, G.; Hu, A.; Zampieri, F.; and De Choudhury, M. 2019b. Discovering alternative treatments for opioid use recovery using social media. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15.

[Chancellor et al.2019c] Chancellor, S.; Nitzburg, G.; Hu, A.; Zampieri, F.; and De Choudhury, M. 2019c. Discovering Alternative Treatments for Opioid Use Recovery Using Social Media. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. Glasgow Scotland Uk: ACM.

[Chancellor, Hu, and De Choudhury2018] Chancellor, S.; Hu, A.; and De Choudhury, M. 2018. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proceedings of the 2018 CHI Conference on Human Factors* *in Computing Systems*, CHI '18, 1–14. Montreal QC, Canada: Association for Computing Machinery.

[Chancellor, Mitra, and De Choudhury2016] Chancellor, S.; Mitra, T.; and De Choudhury, M. 2016. Recovery Amid Pro-Anorexia: Analysis of Recovery in Social Media. 2111–2123. ACM Press.

[Chandrasekharan et al.2017] Chandrasekharan, E.; Samory, M.; Srinivasan, A.; and Gilbert, E. 2017. The Bag of Communities: Identifying Abusive Behavior Online with Preexisting Internet Data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3175–3187. Denver Colorado USA: ACM.

[Chandrasekharan et al.2018] Chandrasekharan, E.; Samory, M.; Jhaver, S.; Charvat, H.; Bruckman, A.; Lampe, C.; Eisenstein, J.; and Gilbert, E. 2018. The Internet's Hidden Rules: An Empirical Study of Reddit Norm Violations at Micro, Meso, and Macro Scales. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW):1–25.

[Chandrasekharan et al.2019] Chandrasekharan, E.; Gandhi, C.; Mustelier, M. W.; and Gilbert, E. 2019. Crossmod: A Cross-Community Learning-based System to Assist Reddit Moderators. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):174:1–174:30.

[Chen et al.2021] Chen, M.; Tworek, J.; Jun, H.; Yuan, Q.; Ponde, H.; Kaplan, J.; Edwards, H.; Burda, Y.; Joseph, N.; Brockman, G.; et al. 2021. Evaluating large language models trained on code. *arXiv preprint arXiv:2107.03374*.

[Cheng and Bernstein2014] Cheng, J., and Bernstein, M. 2014. Catalyst: Triggering collective action with thresholds. 1211–1221. ACM Press.

[Cheng, Danescu-Niculescu-Mizil, and Leskovec2015] Cheng, J.; Danescu-Niculescu-Mizil, C.; and Leskovec, J. 2015. Antisocial Behavior in Online Discussion Communities. 10.

[Choi, Jeon, and Kim2019] Choi, J. P.; Jeon, D.-S.; and Kim, B.-C. 2019. Privacy and personal data collection with information externalities. *Journal of Public Economics* 173:113–124.

[Choudhury and Kiciman2017] Choudhury, M. D., and Kiciman, E. 2017. The Language of Social Support in Social Media and Its Effect on Suicidal Ideation Risk. *Proceedings of the International AAAI Conference on Web and Social Media* (1).

[Commission2017] Commission, E. 2017. Fair Taxation of the Digital Economy. https://ec.europa.eu/taxation_customs/business/company-tax/fair-taxation-digital-economy_en.

[Cong et al.2022] Cong, Z.; Luo, X.; Pei, J.; Zhu, F.; and Zhang, Y. 2022. Data pricing in machine learning pipelines. *Knowledge and Information Systems* 64(6):1417–1455.

[Contractor et al.2022] Contractor, D.; McDuff, D.; Haines, J. K.; Lee, J.; Hines, C.; Hecht, B.; Vincent, N.; and Li, H. 2022. Behavioral use licensing for responsible ai. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, FAccT '22, 778–788. New York, NY, USA: Association for Computing Machinery.

[Crabtree, Rodden, and Benford2005] Crabtree, A.; Rodden, T.; and Benford, S. 2005. Moving with the times: It research and the boundaries of cscw. *Computer Supported Cooperative Work (CSCW)* 14(3):217–251.

[Cummings2006] Cummings, M. L. 2006. Integrating ethics in design through the value-sensitive design approach. *Science and Engineering Ethics* 12(4):701–715.

[Dahl1957] Dahl, R. A. 1957. The concept of power. *Behavioral science* 2(3):201–215.

[Das et al.2021] Das, S.; Edwards, W. K.; Kennedy-Mayo, D.; Swire, P.; and Wu, Y. 2021. Privacy for the people? exploring collective action as a mechanism to shift power to consumers in end-user privacy. *IEEE Security & Privacy* 19(5):66–70.

[De Choudhury et al.2016] De Choudhury, M.; Jhaver, S.; Sugar, B.; and Weber, I. 2016. Social Media Participation in an Activist Movement for Racial Equality. In *Tenth International AAAI Conference on Web and Social Media*.

[de Zúñiga, Copeland, and Bimber2014] de Zúñiga, H. G.; Copeland, L.; and Bimber, B. 2014. Political consumerism: Civic engagement and the social media connection. *New Media & Society* 16(3):488–506.

[Dee2018] 2018. Deep Fakes: A Looming Crisis for National Security, Democracy and Privacy? https://www.lawfareblog.com/deep-fakes-looming-crisis-national-security-democracy-and-privacy.

[Deng et al.2009] Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, 248–255. Ieee.

[D'ignazio and Klein2020] D'ignazio, C., and Klein, L. F. 2020. *Data feminism.* MIT press.

[Dillahunt, Brooks, and Gulati2015] Dillahunt, T. R.; Brooks, C. A.; and Gulati, S. 2015. Detecting and Visualizing Filter Bubbles in Google and Bing. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '15, 1851–1856. New York, NY, USA: ACM.

[DiMaggio et al.2004] DiMaggio, P.; Hargittai, E.; Celeste, C.; and Shafer, S. 2004. Digital inequality: From unequal access to differentiated use. *Social inequality* 355–400.

[Dimond et al.2013] Dimond, J. P.; Dye, M.; Larose, D.; and Bruckman, A. S. 2013. Hollaback!: The Role of Storytelling Online in a Social Movement Organization. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, 477–490. New York, NY, USA: ACM.

[Dingler2016] Dingler, T. 2016. Cognition-aware Systems As Mobile Personal Assistants. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, UbiComp '16, 1035–1040. New York, NY, USA: ACM.

[Dombrowski, Alvarado Garcia, and Despard2017] Dombrowski, L.; Alvarado Garcia, A.; and Despard, J. 2017. Low-Wage Precarious Workers' Sociotechnical Practices Working

Towards Addressing Wage Theft. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 4585–4598.

[Dosono and Semaan2019a] Dosono, B., and Semaan, B. 2019a. Moderation Practices as Emotional Labor in Sustaining Online Communities: The Case of AAPI Identity Work on Reddit. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–13.

[Dosono and Semaan2019b] Dosono, B., and Semaan, B. 2019b. Moderation Practices as Emotional Labor in Sustaining Online Communities: The Case of AAPI Identity Work on Reddit. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, 1–13. Glasgow, Scotland Uk: Association for Computing Machinery.

[Dragicevic2016] Dragicevic, P. 2016. Fair statistical communication in HCI. In *Modern Statistical Methods for HCI*. Springer. 291–330.

[Dube et al.2020] Dube, A.; Jacobs, J.; Naidu, S.; and Suri, S. 2020. Monopsony in Online Labor Markets. *American Economic Review: Insights* 2(1):33–46.

[Epstein et al.2015] Epstein, D. A.; Jacobson, B. H.; Bales, E.; McDonald, D. W.; and Munson, S. A. 2015. From "Nobody Cares" to "Way to Go!": A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, 1622–1636. New York, NY, USA: ACM.

[Epstein et al.2017] Epstein, R.; Robertson, R. E.; Lazer, D.; and Wilson, C. 2017. Suppressing the Search Engine Manipulation Effect (SEME). *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW):1–22.

[Erickson and Kellogg2000] Erickson, T., and Kellogg, W. A. 2000. Social translucence: An approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction* 7(1):59–83.

[Ernala et al.2020] Ernala, S. K.; Burke, M.; Leavitt, A.; and Ellison, N. B. 2020. How Well Do People Report Time Spent on Facebook?: An Evaluation of Established Survey Questions with Recommendations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. Honolulu HI USA: ACM.

[Facebook2021] Facebook. 2021. Community Standards Enforcement Report, Fourth Quarter 2020.

[Fan and Zhang2020a] Fan, J., and Zhang, A. X. 2020a. Digital Juries: A Civics-Oriented Approach to Platform Governance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–14.

[Fan and Zhang2020b] Fan, J., and Zhang, A. X. 2020b. Digital Juries: A Civics-Oriented Approach to Platform Governance. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, 1–14. New York, NY, USA: Association for Computing Machinery.

[Farrell et al.2019] Farrell, T.; Fernandez, M.; Novotny, J.; and Alani, H. 2019. Exploring Misogyny across the Manosphere in Reddit. In *Proceedings of the 10th ACM Conference on Web Science*, WebSci '19, 87–96. Boston, Massachusetts, USA: Association for Computing Machinery.

[Felici, Koulouris, and Pearson2013] Felici, M.; Koulouris, T.; and Pearson, S. 2013. Accountability for Data Governance in Cloud Ecosystems. In *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, volume 2, 327–332.

[Feygin et al.2021] Feygin, Y.; Li, H.; Lala, C.; Hecht, B.; Vincent, N.; Scarcella, L.; and Prewitt, M. 2021. A data dividend that works: steps toward building an equitable data economy.

[Fiesler and Dym2020] Fiesler, C., and Dym, B. 2020. Moving across lands: Online platform migration in fandom communities. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW1):1–25.

[Fiesler and Proferes2018] Fiesler, C., and Proferes, N. 2018. "participant" perceptions of twitter research ethics. *Social Media+ Society* 4(1):2056305118763366.

[Fiesler et al.2018] Fiesler, C.; Jiang, J. A.; McCann, J.; Frye, K.; and Brubaker, J. R. 2018. Reddit rules! Characterizing an ecosystem of governance. In *Proceedings of the AAAI International Conference on Web and Social Media*, 72–81. AAAI.

[Flores-Saviaga, Keegan, and Savage2018] Flores-Saviaga, C. I.; Keegan, B. C.; and Savage, S. 2018. Mobilizing the Trump Train: Understanding Collective Action in a Political

Trolling Community. In *Twelfth International AAAI Conference on Web and Social Media*.

[Flynn2018] Flynn, K. 2018. Facebook's making more money per user in North America than ever before.

[Foong and Gerber2021] Foong, E., and Gerber, E. 2021. Understanding Gender Differences in Pricing Strategies in Online Labor Marketplaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery.

[Foong et al.2018] Foong, E.; Vincent, N.; Hecht, B.; and Gerber, E. M. 2018. Women (Still) Ask For Less: Gender Differences in Hourly Rate in an Online Labor Marketplace. *Proc. ACM Hum.-Comput. Interact.* 2(CSCW).

[Fox et al.2020] Fox, S. E.; Khovanskaya, V.; Crivellaro, C.; Salehi, N.; Dombrowski, L.; Kulkarni, C.; Irani, L.; and Forlizzi, J. 2020. Worker-Centered Design: Expanding HCI Methods for Supporting Labor. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–8. Honolulu HI USA: ACM.

[Fox, Sobel, and Rosner2019] Fox, S. E.; Sobel, K.; and Rosner, D. K. 2019. Managerial Visions: Stories of Upgrading and Maintaining the Public Restroom with IoT. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–15.

[Frauenberger et al.2017] Frauenberger, C.; Bruckman, A. S.; Munteanu, C.; Densmore, M.; and Waycott, J. 2017. Research Ethics in HCI: A Town Hall Meeting. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, 1295–1299. New York, NY, USA: ACM.

[Frauenberger, Rauhala, and Fitzpatrick2017] Frauenberger, C.; Rauhala, M.; and Fitzpatrick, G. 2017. In-Action Ethics. *Interacting with Computers* 29(2):220–236.

[Friedman and Kahn2003] Friedman, B., and Kahn, Jr., P. H. 2003. Human values, ethics, and design. In Jacko, J. A., and Sears, A., eds., *The Human-computer Interaction Handbook*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc. 1177–1201.

[Friedman and Nissenbaum1996] Friedman, B., and Nissenbaum, H. 1996. Bias in Computer Systems. *ACM Trans. Inf. Syst.* 14(3):330–347.

[Friedman, Howe, and Felten2002] Friedman, B.; Howe, D. C.; and Felten, E. 2002. Informed consent in the Mozilla browser: Implementing value-sensitive design. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 10 pp.–.

[Friedman1997] Friedman, B. 1997. *Human Values and the Design of Computer Technology*. Cambridge University Press.

[Friedman2002] Friedman, M. 2002. *Consumer Boycotts: Effecting Change Through the Marketplace and Media*. Routledge.

[G King2011] G King, B. 2011. The Tactical Disruptiveness of Social Movements: Sources of Market and Mediated Disruption in Corporate Boycotts. *Social Problems* 58:491–517.

[Gebru et al.2018] Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Daumé III, H.; and Crawford, K. 2018. Datasheets for datasets. *arXiv preprint arXiv:1803.09010*.

[Gebru et al.2020] Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Daumé III, H.; and Crawford, K. 2020. Datasheets for Datasets. *arXiv:1803.09010 [cs]*.

[Geiger and Halfaker2013a] Geiger, R. S., and Halfaker, A. 2013a. Using Edit Sessions to Measure Participation in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, 861–870. New York, NY, USA: Association for Computing Machinery. event-place: San Antonio, Texas, USA.

[Geiger and Halfaker2013b] Geiger, R. S., and Halfaker, A. 2013b. Using Edit Sessions to Measure Participation in Wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, 861–870. New York, NY, USA: Association for Computing Machinery.

[Geiger and Halfaker2013c] Geiger, R. S., and Halfaker, A. 2013c. Using edit sessions to measure participation in wikipedia. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, CSCW '13, 861–870. San Antonio, Texas, USA: Association for Computing Machinery.

[Geiger et al.2020] Geiger, R. S.; Yu, K.; Yang, Y.; Dai, M.; Qiu, J.; Tang, R.; and Huang, J. 2020. Garbage In, Garbage Out? Do Machine Learning Application Papers in Social Computing Report Where Human-Labeled Training Data Comes From? *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* 325–336.

[Geiger, Howard, and Irani2021] Geiger, R. S.; Howard, D.; and Irani, L. 2021. The labor of maintaining and scaling free and open-source software projects. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW1):1–28.

[Gilbert2020] Gilbert, S. A. 2020. "I run the world's largest historical outreach project and it's on a cesspool of a website." Moderating a Public Scholarship Site on Reddit: A Case Study of r/AskHistorians. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW1):019:1–019:27.

[González Cabañas, Cuevas, and Cuevas2017a] González Cabañas, J.; Cuevas, Á.; and Cuevas, R. 2017a. FDVT: Data Valuation Tool for Facebook Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, 3799–3809. New York, NY, USA: ACM.

[González Cabañas, Cuevas, and Cuevas2017b] González Cabañas, J.; Cuevas, A.; and Cuevas, R. 2017b. Fdvt: Data valuation tool for facebook users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3799–3809.

[gov2019] 2019. A governance framework for algorithmic accountability and transparency.

[Graa] #GrabYourWallet adds Chrome extension to make avoiding Trump brands even easier. https://mic.com/articles/171396/grab-your-wallet-adds-chrome-extension-to-make-avoiding-trump-brands-even-easier.

[Grab] #GrabYourWallet's Anti-Trump Boycott Looks To Expand Its Reach. https://www.npr.org/2017/04/16/523960521/-grabyourwallets-anti-trump-boycott-looks-to-expand-its-reach.

[Granovetter1978] Granovetter, M. 1978. Threshold Models of Collective Behavior. *American Journal of Sociology* 83(6):1420–1443.

[Granville2018] Granville, K. 2018. Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. *N. Y. Times*.

[Gray and Suri2019] Gray, M. L., and Suri, S. 2019. *Ghost work: How to stop Silicon Valley from building a new global underclass*. Eamon Dolan Books.

[Greenfield, Frier, and Brody2018] Greenfield, R.; Frier, S.; and Brody, B. 2018. NAACP Seeks Week-Long Facebook Boycott Over Racial Targeting. *Bloomberg.com*.

[Griffith2017a] Griffith, E. 2017a. Techies Still Think They're the Good Guys. They're Not. — Backchannel. *Wired*.

[Griffith2017b] Griffith, E. 2017b. Techies Still Think They're the Good Guys. They're Not. | Backchannel. *Wired*.

[Griffiths et al.2021] Griffiths, K.; Blain, J.; Vajdic, C.; and Jorm, L. 2021. Indigenous and tribal peoples data governance in health research: A systematic review. *International Journal of Environmental Research and Public Health* 18(19).

[Gro] Groups. https://www.splcenter.org/fighting-hate/extremist-files/groups.

[Grossklags and Acquisti2007] Grossklags, J., and Acquisti, A. 2007. When 25 Cents is Too Much: An Experiment on Willingness-To-Sell and Willingness-To-Protect Personal Information. In *WEIS*.

[Grudin1988] Grudin, J. 1988. Why CSCW applications fail: Problems in the design and evaluationof organizational interfaces. In *Proceedings of the 1988 ACM Conference on Computer-supported Cooperative Work*, CSCW '88, 85–93. New York, NY, USA: Association for Computing Machinery.

[Guha and Birnholtz2013] Guha, S., and Birnholtz, J. 2013. Can you see me now?: Location, visibility and the management of impressions on foursquare. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*, 183–192. ACM.

[Guha and Wicker2015] Guha, S., and Wicker, S. B. 2015. Do Birds of a Feather Watch Each Other?: Homophily and Social Surveillance in Location Based Social Networks. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, 1010–1020. New York, NY, USA: ACM.

[Guha, Baumer, and Gay2018] Guha, S.; Baumer, E. P.; and Gay, G. K. 2018. Regrets, I'Ve Had a Few: When Regretful Experiences Do (and Don'T) Compel Users to Leave Facebook. In *Proceedings of the 2018 ACM Conference on Supporting Groupwork*, GROUP '18, 166–177. New York, NY, USA: ACM.

[Hara et al.2018a] Hara, K.; Adams, A.; Milland, K.; Savage, S.; Callison-Burch, C.; and Bigham, J. P. 2018a. A data-driven analysis of workers' earnings on amazon mechanical turk. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, 1–14.

[Hara et al.2018b] Hara, K.; Adams, A.; Milland, K.; Savage, S.; Callison-Burch, C.; and Bigham, J. P. 2018b. A Data-Driven Analysis of Workers' Earnings on Amazon Mechanical Turk. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–14.

[Hardy and Leiba-O'Sullivan1998] Hardy, C., and Leiba-O'Sullivan, S. 1998. The power behind empowerment: Implications for research and practice. *Human relations* 51(4):451–483.

[Hargittai2007] Hargittai, E. 2007. Whose space? Differences among users and non-users of social network sites. *Journal of computer-mediated communication* 13(1):276–297.

[Harper and Konstan2015] Harper, F. M., and Konstan, J. A. 2015. The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)* 5(4):1–19.

[Hawkins2010] Hawkins, R. A. 2010. Boycotts, buycotts and consumer activism in a global context: An overview. *Management & Organizational History* 5(2):123–143.

[Hayes2014] Hayes, G. R. 2014. Knowing by Doing: Action Research as an Approach to HCI. In Olson, J. S., and Kellogg, W. A., eds., *Ways of Knowing in HCI*. New York, NY: Springer New York. 49–68.

[Heald, Erickson, and Kretschmer2015] Heald, P.; Erickson, K.; and Kretschmer, M. 2015. The Valuation of Unprotected Works: A Case Study of Public Domain Images on Wikipedia. *Harvard Journal of Law & Technology* 29:1.

[Hecht et al.2018] Hecht, B.; Wilcox, L.; Bigham, J.; Schöning, J.; Hoque, E.; Ernst, J.; Bisk, Y.; De Russis, L.; Yarosh, L.; Anjum, B.; Contractor, D.; and Wu, C. 2018. It's Time to Do Something: Mitigating the Negative Impacts of Computing Through a Change to the Peer Review Process.

[Hecht] Hecht, B. B. It's time to do something: Mitigating the negative impacts of computing through a change to the peer review process march 29, 2018— in fca discussions.

[Herndon2019] Herndon, A. W. 2019. Elizabeth Warren Proposes Breaking Up Tech Giants Like Amazon and Facebook. *The New York Times*.

[Hine et al.2017] Hine, G. E.; Onaolapo, J.; Cristofaro, E. D.; Kourtellis, N.; Leontiadis, I.; Samaras, R.; Stringhini, G.; and Blackburn, J. 2017. Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web. In *Eleventh International AAAI Conference on Web and Social Media*.

[Hond and De Bakker2007] Hond, F. D., and De Bakker, F. G. A. 2007. Ideologically Motivated Activism: How Activist Groups Influence Corporate Social Change Activities. *The Academy of Management Review* 32(3):901–924.

[Howe and Nissenbaum2017a] Howe, D. C., and Nissenbaum, H. 2017a. Engineering privacy and protest: A case study of AdNauseam. In *CEUR Workshop Proceedings*, volume 1873, 57–64.

[Howe and Nissenbaum2017b] Howe, D. C., and Nissenbaum, H. 2017b. Engineering privacy and protest: A case study of adnauseam. In *IWPE@ SP*, 57–64.

[ICP] Inter-university consortium for political and social research.

[Irani and Silberman2013a] Irani, L. C., and Silberman, M. S. 2013a. Turkopticon: Interrupting worker invisibility in amazon mechanical turk. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 611–620.

[Irani and Silberman2013b] Irani, L. C., and Silberman, M. S. 2013b. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 611–620.

[Isaac2018] Isaac, M. 2018. What You Need to Know About #DeleteUber. *The New York Times*.

[Janssen et al.2020] Janssen, M.; Brous, P.; Estevez, E.; Barbosa, L. S.; and Janowski, T. 2020. Data governance: Organizing data for trustworthy Artificial Intelligence. *Government Information Quarterly* 37(3):101493.

[Jhaver et al.2019a] Jhaver, S.; Appling, D. S.; Gilbert, E.; and Bruckman, A. 2019a. "Did You Suspect the Post Would be Removed?": Understanding User Reactions to Content Removals on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):192:1–192:33.

[Jhaver et al.2019b] Jhaver, S.; Birman, I.; Gilbert, E.; and Bruckman, A. 2019b. Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator. *ACM Transactions on Computer-Human Interaction* 26(5):1–35.

[Jhaver, Bruckman, and Gilbert2019a] Jhaver, S.; Bruckman, A.; and Gilbert, E. 2019a. Does Transparency in Moderation Really Matter? User Behavior After Content Removal Explanations on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):150:1–150:27.

[Jhaver, Bruckman, and Gilbert2019b] Jhaver, S.; Bruckman, A.; and Gilbert, E. 2019b. Does Transparency in Moderation Really Matter? User Behavior After Content Removal Explanations on Reddit. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):150:1–150:27.

[Jhaver, Chan, and Bruckman2018] Jhaver, S.; Chan, L.; and Bruckman, A. 2018. The view from the other side: The border between controversial speech and harassment on Kotaku in Action. *First Monday* 23(2).

[Johnson et al.2016a] Johnson, I. L.; Lin, Y.; Li, T. J.-J.; Hall, A.; Halfaker, A.; Schöning, J.; and Hecht, B. 2016a. Not at Home on the Range: Peer Production and the Urban/Rural Divide. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 13–25. New York, NY, USA: ACM.

[Johnson et al.2016b] Johnson, I. L.; Lin, Y.; Li, T. J.-J.; Hall, A.; Halfaker, A.; Schöning, J.; and Hecht, B. 2016b. Not at home on the range: Peer production and the urban/rural divide. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 13–25. New York, NY, USA: Association for Computing Machinery.

[Jones and Tonetti2020] Jones, C. I., and Tonetti, C. 2020. Nonrivalry and the economics of data. *American Economic Review* 110(9):2819–58.

[Kasperkevic2018] Kasperkevic, J. 2018. Amazon faces boycott ahead of holidays as public discontent grows. *the Guardian*.

[Khovanskaya et al.2019a] Khovanskaya, V.; Dombrowski, L.; Rzeszotarski, J.; and Sengers, P. 2019a. The tools of management: Adapting historical union tactics to platform-mediated labor. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):1–22.

[Khovanskaya et al.2019b] Khovanskaya, V.; Dombrowski, L.; Rzeszotarski, J.; and Sengers, P. 2019b. The Tools of Management: Adapting Historical Union Tactics to Platform-Mediated Labor. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):1–22.

[Kiene and Hill2020] Kiene, C., and Hill, B. M. 2020. Who Uses Bots? A Statistical Analysis of Bot Usage in Moderation Teams. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, 1–8. New York, NY, USA: Association for Computing Machinery.

[Kim and Luca2019] Kim, H., and Luca, M. 2019. Product Quality and Entering Through Tying: Experimental Evidence. *Management Science* 65(2):596–603.

[Kim, Hullman, and Agrawala2016] Kim, Y.-S.; Hullman, J.; and Agrawala, M. 2016. Generating Personalized Spatial Analogies for Distances and Areas. 38–48. ACM Press.

[King and McDonnell2015] King, B. G., and McDonnell, M. H. 2015. Good firms, good targets: The relationship among corporate social responsibility, reputation, and activist targeting. *Corporate Social Responsibility in a Globalizing World* 430–454.

[Kittur et al.2013] Kittur, A.; Nickerson, J. V.; Bernstein, M.; Gerber, E.; Shaw, A.; Zimmerman, J.; Lease, M.; and Horton, J. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 1301–1318.

[Klein, Smith, and John2004] Klein, J. G.; Smith, N. C.; and John, A. 2004. Why We Boycott: Consumer Motivations for Boycott Participation. *Journal of Marketing* 68(3):92–109.

[Koos2012] Koos, S. 2012. What drives political consumption in Europe? A multi-level analysis on individual characteristics, opportunity structures and globalization. *Acta Sociologica* 55(1):37–57.

[Kramer, Guillory, and Hancock2014] Kramer, A. D. I.; Guillory, J. E.; and Hancock, J. T. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences* 201320040.

[Kraut et al.2012] Kraut, R. E.; Resnick, P.; Kiesler, S.; Burke, M.; Chen, Y.; Kittur, N.; Konstan, J.; Ren, Y.; and Riedl, J. 2012. *Building Successful Online Communities: Evidence-Based Social Design*. MIT Press.

[Kriplean, Beschastnikh, and McDonald2008] Kriplean, T.; Beschastnikh, I.; and McDonald, D. W. 2008. Articulations of Wikiwork: Uncovering Valued Work in Wikipedia through Barnstars. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*, CSCW '08, 47–56. New York, NY, USA: Association for Computing Machinery. event-place: San Diego, CA, USA.

[Lampe, Vitak, and Ellison2013] Lampe, C.; Vitak, J.; and Ellison, N. 2013. Users and nonusers: Interactions between levels of adoption and social capital. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, 809–820. ACM.

[Lanier and Weyl2018a] Lanier, J., and Weyl, E. G. 2018a. A Blueprint for a Better Digital Society. 18.

[Lanier and Weyl2018b] Lanier, J., and Weyl, E. G. 2018b. A blueprint for a better digital society. *Harvard Business Review* 26.

[Le Dantec, Poole, and Wyche2009] Le Dantec, C. A.; Poole, E. S.; and Wyche, S. P. 2009. Values as Lived Experience: Evolving Value Sensitive Design in Support of Value

Discovery. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1141–1150.

[Lee et al.2019] Lee, M. K.; Kusbit, D.; Kahng, A.; Kim, J. T.; Yuan, X.; Chan, A.; See, D.; Noothigattu, R.; Lee, S.; Psomas, A.; et al. 2019. Webuildai: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):1–35.

[Li and Hecht2021] Li, H., and Hecht, B. 2021. 3 stars on yelp, 4 stars on google maps: A cross-platform examination of restaurant ratings. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW3):1–25.

[Li et al.2018a] Li, H.; Alarcon, B.; Espinosa, S. M.; and Hecht, B. 2018a. Out of Site: Empowering a New Approach to Online Boycotts. *Proceedings of the 2018 Computer-Supported Cooperative Work and Social Computing (CSCW'2018 / PACM).*

[Li et al.2018b] Li, H.; Alarcon, B.; Milkes Espinosa, S.; and Hecht, B. 2018b. Out of Site: Empowering a New Approach to Online Boycotts. *Proc. ACM Hum.-Comput. Interact.* 2(CSCW):106:1–106:28.

[Li et al.2019a] Li, H.; Vincent, N.; Tsai, J.; Kaye, J.; and Hecht, B. 2019a. How Do People Change Their Technology Use in Protest? Understanding. *Proc. ACM Hum.-Comput. Interact.* 3(CSCW).

[Li et al.2019b] Li, H.; Vincent, N.; Tsai, J.; Kaye, J.; and Hecht, B. 2019b. How do people change their technology use in protest?: Understanding "protest users". *Proceedings of the ACM on Human-Computer Interaction* 3(CSCW):87.

[Li, Hecht, and Chancellor2022] Li, H.; Hecht, B.; and Chancellor, S. 2022. All that's happening behind the scenes: Putting the spotlight on volunteer moderator labor in reddit. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, 584–595.

[Lin et al.2017a] Lin, Z.; Salehi, N.; Yao, B.; Chen, Y.; and Bernstein, M. 2017a. Better When It Was Smaller? Community Content and Behavior After Massive Growth. *Proceedings of the International AAAI Conference on Web and Social Media* 11(1).

[Lin et al.2017b] Lin, Z.; Salehi, N.; Yao, B.; Chen, Y.; and Bernstein, M. S. 2017b. Better When It Was Smaller? Community Content and Behavior After Massive Growth. 10.

[Ling et al.2005] Ling, K.; Beenen, G.; Ludford, P.; Wang, X.; Chang, K.; Li, X.; Cosley, D.; Frankowski, D.; Terveen, L.; Rashid, A. M.; Resnick, P.; and Kraut, R. 2005. Using Social Psychology to Motivate Contributions to Online Communities. *Journal of Computer-Mediated Communication* 10(4):00–00.

[Lo2018] Lo, C. 2018. *When All You Have Is a Banhammer: The Social and Communicative Work of Volunteer Moderators.* Thesis.

[Lokot and Diakopoulos2016] Lokot, T., and Diakopoulos, N. 2016. News Bots: Automating news and information dissemination on Twitter. *Digital Journalism* 4(6):682–699.

[Lotz2019] Lotz, A. 2019. *'Big Tech' Isn't One Big Monopoly – It's 5 Companies All in Different Businesses*.

[Lu et al.2014] Lu, C.-T.; Xie, S.; Kong, X.; and Yu, P. S. 2014. Inferring the Impacts of Social Media on Crowdfunding. In *Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, WSDM '14, 573–582. New York, NY, USA: ACM.

[Manjoo2018] Manjoo, F. 2018. Stumbles? What Stumbles? Big Tech Is as Strong as Ever. *The New York Times*.

[Martínez2019] Martínez, A. G. 2019. Facebook Is Not a Monopoly, but It Should Be Broken Up. *Wired*.

[Masood et al.2018] Masood, R.; Vatsalan, D.; Ikram, M.; and Kaafar, M. A. 2018. Incognito: A Method for Obfuscating Web Data. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 267–276. Lyon, France: ACM Press.

[Mathur et al.2018a] Mathur, A.; Vitak, J.; Narayanan, A.; and Chetty, M. 2018a. Characterizing the Use of Browser-Based Blocking Extensions To Prevent Online Tracking. In *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*, 103–116.

[Mathur et al.2018b] Mathur, A.; Vitak, J.; Narayanan, A.; and Chetty, M. 2018b. Characterizing the use of browser-based blocking extensions to prevent online tracking. In *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*, 103–116.

[Matias2016a] Matias, J. N. 2016a. The civic labor of online moderators. *The Platform Society* 10.

[Matias2016b] Matias, J. N. 2016b. Going dark: Social factors in collective action against platform operators in the Reddit blackout. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*, 1138–1151. New York, NY: ACM.

[Matias2016c] Matias, J. N. 2016c. Going dark: Social factors in collective action against platform operators in the reddit blackout. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 1138–1151.

[Matias2016d] Matias, J. N. 2016d. Going Dark: Social Factors in Collective Action Against Platform Operators in the Reddit Blackout. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, 1138–1151. New York, NY, USA: ACM.

[Matias2019a] Matias, J. N. 2019a. The civic labor of volunteer moderators online. *Social Media+ Society* 5(2):2056305119836778.

[Matias2019b] Matias, J. N. 2019b. The Civic Labor of Volunteer Moderators Online. *Social Media + Society* 5(2):2056305119836778.

[McDonnell, King, and Soule2015] McDonnell, M.-H.; King, B. G.; and Soule, S. A. 2015. A Dynamic Process Model of Private Politics: Activist Targeting and Corporate Receptivity to Social Challenges. *American Sociological Review* 80(3):654–678.

[McMahon, Johnson, and Hecht2017a] McMahon, C.; Johnson, I. L.; and Hecht, B. 2017a. The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies. In *ICWSM*, 142–151.

[McMahon, Johnson, and Hecht2017b] McMahon, C.; Johnson, I. L.; and Hecht, B. 2017b. The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies. In *ICWSM*, 142–151.

[Micheletti, Follesdal, and Stolle] Micheletti, M.; Follesdal, A.; and Stolle, D. Politics, Products, and Markets: Exploring Political Consumerism Past and Present. *Economic Geography* 84(1):123–125.

[Møller et al.2020] Møller, N. H.; Bossen, C.; Pine, K. H.; Nielsen, T. R.; and Neff, G. 2020. Who does the work of data? *Interactions* 27(3):52–55.

[mon] The montréal declaration responsible ai.

[Muller2003a] Muller, M. J. 2003a. The Human-computer Interaction Handbook. Hillsdale, NJ, USA: L. Erlbaum Associates Inc. 1051–1068.

[Muller2003b] Muller, M. J. 2003b. Participatory design: The third space in HCI. In Jacko, J. A., and Sears, A., eds., *The Human-computer Interaction Handbook*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc. 1051–1068.

[Munson, Lee, and Resnick2013] Munson, S. A.; Lee, S. Y.; and Resnick, P. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget. In *Proceedings of the 7th International Conference on Weblogs and Social Media, ICWSM 2013*.

[Myers West2018] Myers West, S. 2018. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. *New Media & Society* 20(11):4366–4383.

[Newell et al.2016] Newell, E.; Jurgens, D.; Saleem, H. M.; Vala, H.; Sassine, J.; Armstrong, C.; and Ruths, D. 2016. User migration in online social networks: A case study on reddit during a period of community unrest. In *ICWSM*, 279–288.

[Newholm and Shaw2007] Newholm, T., and Shaw, D. 2007. Studying the ethical consumer: A review of research. *Journal of Consumer Behaviour* 6(5):253–270.

[Newholm2005] Newholm, T. 2005. Case Studying Ethical Consumers' Projects and Strategies. In *The Ethical Consumer*. London: SAGE Publications Ltd. 107–124.

[Newman and Bartels2011] Newman, B. J., and Bartels, B. L. 2011. Politics at the Checkout Line: Explaining Political Consumerism in the United States. *Political Research Quarterly* 64(4):803–817.

[Nguyen et al.2014] Nguyen, T. T.; Hui, P.-M.; Harper, F. M.; Terveen, L.; and Konstan, J. A. 2014. Exploring the Filter Bubble: The Effect of Using Recommender Systems on Content Diversity. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, 677–686. New York, NY, USA: ACM.

[Nielsen and Molich1990] Nielsen, J., and Molich, R. 1990. Heuristic Evaluation of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '90, 249–256. New York, NY, USA: ACM.

[Nielsen1994a] Nielsen, J. 1994a. Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier. In Bias, R. G., and Mayhew, D. J., eds., *Cost-Justifying Usability*. Orlando, FL, USA: Academic Press, Inc. 245–272.

[Nielsen1994b] Nielsen, J. 1994b. *Usability Engineering*. Elsevier.

[NYU] Nyu ad observatory by nyu cybersecurity for democracy.

[Oliver, Marwell, and Teixeira1985] Oliver, P.; Marwell, G.; and Teixeira, R. 1985. A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action. *American Journal of Sociology* 91(3):522–556.

[OLSON2009] OLSON, M. 2009. *THE LOGIC OF COLLECTIVE ACTION*. Harvard University Press.

[osl] About open source licenses.

[Ott et al.2011] Ott, M.; Choi, Y.; Cardie, C.; and Hancock, J. T. 2011. Finding deceptive opinion spam by any stretch of the imagination. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, 309–319. Portland, Oregon: Association for Computational Linguistics.

[Overdorf et al.2018] Overdorf, R.; Kulynych, B.; Balsa, E.; Troncoso, C.; and Gürses, S. 2018. POTs: Protective Optimization Technologies. *arXiv preprint arXiv:1806.02711*.

[Parikh2018] Parikh, T. 2018. Mitigating the Negative Implications of Computing: Making Space for Debate.

[Pariser2012] Pariser, E. 2012. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books.

[Patel2020] Patel, S. 2020. Reddit Claims 52 Million Daily Users, Revealing a Key Figure for Social-Media Platforms. *Wall Street Journal*.

[Peck2019] Peck, R. 2019. The Punishing Ecstasy of Being a Reddit Moderator. *Wired*.

[Pei et al.2021] Pei, J.; Zhu, F.; Cong, Z.; Luo, X.; Liu, H.; and Mu, X. 2021. Data Pricing and Data Asset Governance in the AI Era. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, KDD '21, 4058–4059. New York, NY, USA: Association for Computing Machinery.

[Pei2020] Pei, J. 2020. A Survey on Data Pricing: From Economics to Data Science. *IEEE Transactions on Knowledge and Data Engineering* 1–1.

[Peng, Mathur, and Narayanan2021] Peng, K.; Mathur, A.; and Narayanan, A. 2021. Mitigating dataset harms requires stewardship: Lessons from 1000 papers. *arXiv preprint arXiv:2108.02922*.

[Pentland and Hardjono2020] Pentland, A., and Hardjono, T. 2020. 2. data cooperatives. In *Building the New Economy*. 0 edition. https://wip.mitpress.mit.edu/pub/pnxgvubq.

[Piccardi et al.2020] Piccardi, T.; Redi, M.; Colavizza, G.; and West, R. 2020. Quantifying engagement with citations on wikipedia. In *Proceedings of The Web Conference 2020*, 2365–2376.

[Piccardi et al.2021a] Piccardi, T.; Redi, M.; Colavizza, G.; and West, R. 2021a. On the value of wikipedia as a gateway to the web. In *Proceedings of the Web Conference 2021*, 249–260.

[Piccardi et al.2021b] Piccardi, T.; Redi, M.; Colavizza, G.; and West, R. 2021b. On the Value of Wikipedia as a Gateway to the Web. In *Proceedings of the Web Conference 2021*. New York, NY, USA: Association for Computing Machinery. 249–260.

[Por2018] 2018. Art. 20 GDPR – Right to data portability | General Data Protection Regulation (GDPR).

[Posner and Weyl2018] Posner, E. A., and Weyl, E. G. 2018. *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press.

[Posner and Weyl2019] Posner, E. A., and Weyl, E. G. 2019. *Radical Markets*. Princeton University Press.

[Post2020] Post, W. 2020. Burnout, splinter factions and deleted posts: Unpaid online moderators struggle to manage divided communities. *Washington Post*.

[Postigo2009] Postigo, H. 2009. America Online volunteers: Lessons from an early co-production community. *International Journal of Cultural Studies* 12(5):451–469.

[Postigo2016] Postigo, H. 2016. The socio-technical architecture of digital labor: Converting play into youtube money. *New media & society* 18(2):332–349.

[Pri] Prime Day: Activists protest against Amazon in cities across US — Technology — The Guardian. https://www.theguardian.com/technology/2019/jul/15/prime-day-activists-plan-protests-in-us-cities-and-a-boycott-of-e-commerce-giant.

[Proferes et al.2021] Proferes, N.; Jones, N.; Gilbert, S.; Fiesler, C.; and Zimmer, M. 2021. Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics. *Social Media + Society* 7(2):20563051211019004.

[Quinn and Bederson2011] Quinn, A. J., and Bederson, B. B. 2011. Human Computation: A Survey and Taxonomy of a Growing Field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1403–1412.

[Radford et al.2019] Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog* 1(8):9.

[Rae et al.2021] Rae, J. W.; Borgeaud, S.; Cai, T.; Millican, K.; Hoffmann, J.; Song, F.; Aslanides, J.; Henderson, S.; Ring, R.; Young, S.; Rutherford, E.; Hennigan, T.; Menick, J.; Cassirer, A.; Powell, R.; Driessche, G. v. d.; Hendricks, L. A.; Rauh, M.; Huang, P.-S.; Glaese, A.; Welbl, J.; Dathathri, S.; Huang, S.; Uesato, J.; Mellor, J.; Higgins, I.; Creswell, A.; McAleese, N.; Wu, A.; Elsen, E.; Jayakumar, S.; Buchatskaya, E.; Budden, D.; Sutherland, E.; Simonyan, K.; Paganini, M.; Sifre, L.; Martens, L.; Li, X. L.; Kuncoro, A.; Nematzadeh, A.; Gribovskaya, E.; Donato, D.; Lazaridou, A.; Mensch, A.;

Lespiau, J.-B.; Tsimpoukelli, M.; Grigorev, N.; Fritz, D.; Sottiaux, T.; Pajarskas, M.; Pohlen, T.; Gong, Z.; Toyama, D.; d'Autume, C. d. M.; Li, Y.; Terzi, T.; Mikulik, V.; Babuschkin, I.; Clark, A.; Casas, D. d. L.; Guy, A.; Jones, C.; Bradbury, J.; Johnson, M.; Hechtman, B.; Weidinger, L.; Gabriel, I.; Isaac, W.; Lockhart, E.; Osindero, S.; Rimell, L.; Dyer, C.; Vinyals, O.; Ayoub, K.; Stanway, J.; Bennett, L.; Hassabis, D.; Kavukcuoglu, K.; and Irving, G. 2021. Scaling Language Models: Methods, Analysis & Insights from Training Gopher.

[Rashid et al.2006] Rashid, A. M.; Ling, K.; Tassone, R. D.; Resnick, P.; Kraut, R.; and Riedl, J. 2006. Motivating Participation by Displaying the Value of Contribution. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, 955–958. New York, NY, USA: ACM.

[Redmiles, Kross, and Mazurek2016] Redmiles, E. M.; Kross, S.; and Mazurek, M. L. 2016. How I Learned to Be Secure: A Census-Representative Survey of Security Advice Sources and Behavior. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS '16, 666–677. New York, NY, USA: ACM.

[Res] Resistbot. https://resist.bot.

[Resnick et al.2013] Resnick, P.; Garrett, R. K.; Kriplean, T.; Munson, S. A.; and Stroud, N. J. 2013. Bursting Your (Filter) Bubble: Strategies for Promoting Diverse Exposure. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*, CSCW '13, 95–100. New York, NY, USA: ACM.

[Roberts2019] Roberts, S. T. 2019. *Behind the Screen*. Yale University Press.

[Rogoff2019] Rogoff, K. 2019. Big tech has too much monopoly power – it's right to take it on. *the Guardian*.

[Ruckenstein and Turunen2019a] Ruckenstein, M., and Turunen, L. L. M. 2019a. Re-humanizing the platform: Content moderators and the logic of care:. *New Media & Society*.

[Ruckenstein and Turunen2019b] Ruckenstein, M., and Turunen, L. L. M. 2019b. Re-humanizing the platform: Content moderators and the logic of care:. *New Media & Society*. Publisher: SAGE PublicationsSage UK: London, England.

[Sadowski, Viljoen, and Whittaker2021] Sadowski, J.; Viljoen, S.; and Whittaker, M. 2021. Everyone should decide how their digital data are used — not just tech companies. *Nature* 595(7866):169–171.

[Sadowski2019] Sadowski, J. 2019. When data is capital: Datafication, accumulation, and extraction. *Big Data & Society* 6(1):2053951718820549.

[Salehi et al.2015] Salehi, N.; Irani, L. C.; Bernstein, M. S.; Alkhatib, A.; Ogbe, E.; and Milland, K. 2015. We are dynamo: Overcoming stalling and friction in collective action for crowd workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 1621–1630.

[Sambasivan et al.2021] Sambasivan, N.; Kapania, S.; Highfill, H.; Akrong, D.; Paritosh, P.; and Aroyo, L. M. 2021. "everyone wants to do the model work, not the data work":

Data cascades in high-stakes ai. In *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15.

[Sannon and Cosley2019a] Sannon, S., and Cosley, D. 2019a. Privacy, Power, and Invisible Labor on Amazon Mechanical Turk. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery. 1–12.

[Sannon and Cosley2019b] Sannon, S., and Cosley, D. 2019b. Privacy, power, and invisible labor on amazon mechanical turk. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–12.

[Sannon, Bazarova, and Cosley2018] Sannon, S.; Bazarova, N. N.; and Cosley, D. 2018. Privacy lies: Understanding how, when, and why people lie to protect their privacy in multiple online contexts. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 52. ACM.

[Satchell and Dourish2009] Satchell, C., and Dourish, P. 2009. Beyond the user: Use and non-use in HCI. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, 9–16. ACM.

[Schmidt2011] Schmidt, K. 2011. The concept of 'work'in cscw. *Computer Supported Cooperative Work (CSCW)* 20(4):341–401.

[Schneider et al.2018] Schneider, H.; Eiband, M.; Ullrich, D.; and Butz, A. 2018. Empowerment in hci-a survey and framework. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–14.

[Schoenebeck2014] Schoenebeck, S. Y. 2014. Giving Up Twitter for Lent: How and Why We Take Breaks from Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, 773–782. New York, NY, USA: ACM.

[Scholz2012] Scholz, T. 2012. *Digital labor: The internet as playground and factory*. Routledge.

[Seering et al.2019] Seering, J.; Wang, T.; Yoon, J.; and Kaufman, G. 2019. Moderator engagement and community development in the age of algorithms. *New Media & Society* 21(7):1417–1443.

[Seering, Kaufman, and Chancellor2020] Seering, J.; Kaufman, G.; and Chancellor, S. 2020. Metaphors in moderation. *New Media & Society* 1461444820964968.

[Seering2020] Seering, J. 2020. Reconsidering Self-Moderation: The Role of Research in Supporting Community-Based Models for Online Content Moderation. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW2):107:1–107:28.

[Selbst et al.2019] Selbst, A. D.; Boyd, D.; Friedler, S. A.; Venkatasubramanian, S.; and Vertesi, J. 2019. Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency*, 59–68.

[Semuels2017] Semuels, A. 2017. Why #DeleteUber and Other Boycotts Matter. *Atlantic*.

[Shao2003] Shao, J. 2003. Impact of the bootstrap on sample surveys. *Statistical Science* 18(2):191–198.

[Shaw and Hargittai2018] Shaw, A., and Hargittai, E. 2018. The Pipeline of Online Participation Inequalities: The Case of Wikipedia Editing. *Journal of Communication* 68(1):143–168.

[Shaw et al.2014] Shaw, A.; Zhang, H.; Monroy-Hernández, A.; Munson, S.; Hill, B. M.; Gerber, E.; Kinnaird, P.; and Minder, P. 2014. Computer supported collective action. *interactions* 21(2):74–77.

[Solon2019] Solon, O. 2019. Facial recognition's 'dirty little secret': Millions of online photos scraped without consent. *NBC News*.

[Sta2017] 2017. Standing up for what's right — Uber Newsroom US. https://www.uber.com/newsroom/standing-up-for-whats-right-3/.

[Star and Strauss1999] Star, S. L., and Strauss, A. 1999. Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work (CSCW)* 8(1):9–30.

[Starbird2013] Starbird, K. 2013. Delivering Patients to Sacré Coeur: Collective Intelligence in Digital Volunteer Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 801–810. New York, NY, USA: ACM.

[Stieger et al.2013] Stieger, S.; Burger, C.; Bohn, M.; and Voracek, M. 2013. Who commits virtual identity suicide? Differences in privacy concerns, internet addiction, and personality between Facebook users and quitters. *Cyberpsychology, Behavior, and Social Networking* 16(9):629–634.

[Stoddard2015] Stoddard, G. 2015. Popularity Dynamics and Intrinsic Quality in Reddit and Hacker News. 10.

[Stolle, Hooghe, and Micheletti2005] Stolle, D.; Hooghe, M.; and Micheletti, M. 2005. Politics in the Supermarket: Political Consumerism as a Form of Political Participation. *International Political Science Review* 26(3):245–269.

[Suchman1995a] Suchman, L. 1995a. Making work visible. *Communications of the ACM* 38(9):56–64.

[Suchman1995b] Suchman, L. 1995b. Making work visible. *Communications of the ACM* 38(9):56–64.

[Terranova2000a] Terranova, T. 2000a. Free labor: Producing culture for the digital economy. *Social text* 18(2):33–58.

[Terranova2000b] Terranova, T. 2000b. Free Labor: Producing Culture for the Digital Economy. *Social Text* 18.

[the]

[To et al.2020] To, A.; Sweeney, W.; Hammer, J.; and Kaufman, G. 2020. "They Just Don't Get It": Towards Social Technologies for Coping with Interpersonal Racism. *Proc.*

ACM Hum.-Comput. Interact. 4(CSCW1). Place: New York, NY, USA Publisher: Association for Computing Machinery.

[Toxtli, Suri, and Savage2021] Toxtli, C.; Suri, S.; and Savage, S. 2021. Quantifying the Invisible Labor in Crowd Work. *arXiv:2110.00169 [cs]*.

[Treem and Leonardi2013] Treem, J. W., and Leonardi, P. M. 2013. Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association. *Annals of the International Communication Association* 36(1):143–189.

[Typ] Types of Discrimination. https://www.eeoc.gov/laws/types/.

[Van Alstyne and Brynjolfsson2005] Van Alstyne, M., and Brynjolfsson, E. 2005. Global Village or Cyber-Balkans? Modeling and Measuring the Integration of Electronic Communities. *Management Science* 51(6):851–868.

[Vardi2018] Vardi, M. Y. 2018. Move Fast and Break Things. *Communications of the ACM* 61(9):7.

[Vasi and Macy2002] Vasi, I. B., and Macy, M. 2002. The Mobilizer's Dilemma: Crisis, Empowerment, and Collective Action. *Social Forces* 81:979–998.

[Veselovsky et al.2022] Veselovsky, V.; Sarkar, D.; Anderson, J.; and Soden, R. 2022. An automated approach to identifying corporate editing. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, 1052–1063.

[Viljoen2020] Viljoen, S. 2020. Democratic data: A relational theory for data governance. *Available at SSRN 3727562*.

[Vincent and Hecht2021] Vincent, N., and Hecht, B. 2021. A deeper investigation of the importance of wikipedia links to search engine results. *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW1):1–15.

[Vincent et al.2021a] Vincent, N.; Li, H.; Tilly, N.; Chancellor, S.; and Hecht, B. 2021a. 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*, 215–227.

[Vincent et al.2021b] Vincent, N.; Li, H.; Tilly, N.; Chancellor, S.; and Hecht, B. 2021b. 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*. New York, NY, USA: Association for Computing Machinery. 215–227.

[Vincent, Hecht, and Sen2019] Vincent, N.; Hecht, B.; and Sen, S. 2019. "Data Strikes": Evaluating the Effectiveness of New Forms of Collective Action Against Technology Platforms. In *Proceedings of The Web Conference 2019*.

[Vincent, Johnson, and Hecht2018] Vincent, N.; Johnson, I.; and Hecht, B. 2018. Examining wikipedia with a broader lens: Quantifying the value of wikipedia's relationships with other large-scale online communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13.

[Whitney et al.2021] Whitney, C. D.; Naval, T.; Quepons, E.; Singh, S.; Rick, S. R.; and Irani, L. 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*, 1–15.

[Why] Why Thousands Of Amazon Workers Are Striking On Prime Day — HuffPost. https://www.huffpost.com/entry/amazon-prime-day-minnesota-strike-germany-protest_n_5d2d52f8e4b02fd71dd95281.

[wik2022] 2022. Wikimedia enterprise announces google and internet archive as its first customers; allows new customers to self sign-up for free trials.

[Wired2014] Wired. 2014. The Laborers Who Keep Dick Pics and Beheadings Out of Your Facebook Feed — WIRED. https://www.wired.com/2014/10/content-moderation/.

[Wom] The woman behind the boycott that is pressuring retailers to dump the Trumps. https://www.washingtonpost.com/news/business/wp/2017/02/13/the-woman-behind-the-boycott-that-is-pressuring-retailers-to-dump-the-trumps/.

[Wright2018] Wright, R. 2018. Why We Can't Let Google Monopolize AI. *Wired*.

[Wyatt, Oudshoorn, and Pinch2003] Wyatt, S. M.; Oudshoorn, N.; and Pinch, T. 2003. Non-users also matter: The construction of users and non-users of the Internet. *Now users matter: The co-construction of users and technology* 67–79.

[Young and Quan-Haase2013] Young, A. L., and Quan-Haase, A. 2013. Privacy Protection Strategies on Facebook. *Information, Communication & Society* 16(4):479–500.

[Zhang et al.2022] Zhang, C. C.; Houtti, M.; Smith, C. E.; Kong, R.; and Terveen, L. 2022. Working for the invisible machines or pumping information into an empty void? an exploration of wikidata contributors' motivations. *Proceedings of the ACM on Human-Computer Interaction* 6(CSCW1):1–21.

[Zhang, Muller, and Wang2020] Zhang, A. X.; Muller, M.; and Wang, D. 2020. How do Data Science Workers Collaborate? Roles, Workflows, and Tools. *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW1):022:1–022:23.

[Zhu et al.2013] Zhu, H.; Zhang, A.; He, J.; Kraut, R. E.; and Kittur, A. 2013. Effects of Peer Feedback on Contribution: A Field Experiment in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, 2253–2262. New York, NY, USA: ACM.

[Zhu et al.2018] Zhu, H.; Yu, B.; Halfaker, A.; and Terveen, L. 2018. Value-sensitive algorithm design: Method, case study, and lessons. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW):1–23.