Original Research Article

BIG DATA & SOCIETY

Big Data & Society
July–December: 1–13

DOI: 10.1177/20539517211035673
journals.sagepub.com/home/bds
SAGE

"The revolution will not be supervised": Consent and open secrets in data science

Coleen Carrigan, Madison W Green and Abibat Rahman-Davies

Abstract

The social impacts of computer technology are often glorified in public discourse, but there is growing concern about its actual effects on society. In this article, we ask: how does "consent" as an analytical framework make visible the social dynamics and power relations in the capture, extraction, and labor of data science knowledge production? We hypothesize that a form of boundary violation in data science workplaces—gender harassment—may correlate with the ways humans' lived experiences are extracted to produce Big Data. The concept of consent offers a useful way to draw comparisons between gender relations in data science and the means by which machines are trained to learn and reason. Inspired by how Big Tech leaders describe unsupervised machine learning, and the co-optation of "revolutionary" rhetoric they use to do so, we introduce a concept we call "techniques of invisibility." Techniques of invisibility are the ways in which an extreme imbalance between exposure and opacity, demarcated along fault lines of power, are fabricated and maintained, closing down the possibility for bidirectional transparency in the production and applications of algorithms. Further, techniques of invisibility, which we group into two categories—epistemic injustice and the Brotherhood—include acts of subjection by powerful actors in data science designed to quell resistance to exploitative relations. These techniques may be useful in making further connections between epistemic violence, sexism, and surveillance, sussing out persistent boundary violations in data science to render the social in data science visible, and open to scrutiny and debate.

Keywords

Feminism, surveillance, gender harassment, data science, consent, privacy

> …The revolution will put you in the driver's seat
> the revolution will not be televised.
>
> —Gil Scott-Heron, 1971

Introduction

Computer technology, long glorified in public discourse, is now under greater scrutiny. In the past 10 years, a sense of urgency has infused debates about Big Tech's negative effects on society, such as surveillance, privacy transgressions, extreme wealth disparities, and labor exploitation. In this paper, we join activists, tech workers, and other scholars, who are also debating the sophisticated mechanisms of surveillance and resisting inequities in data science classrooms and workplaces. Extending earlier scholarship (Adam, 1998), these critiques disrupt the ennobling discourse on the benevolent influence of data science in US society that could have tangible benefits for advancing justice. Our intent is to make visible obscured relations of power using a comparative optics framework of consent to analyze common features between gendered dynamics within local sites of data science knowledge production and the further enclosure of social life within Big Tech's modes of production.

We investigate if power relations in sites of data science knowledge production influence the outputs of these workplaces. Specifically, we consider if gender harassment, a form of boundary violation that consists of verbal, physical, and symbolic behaviors conveying hostility toward women and non-binary people, correlates with the processes by which humans' lived experiences are extracted to train machines. We first examine consent and refusal as they relate to civility in the workplace and privacy in the

Social Sciences Department, California State Polytechnic University, USA

Corresponding author:
Coleen Carrigan, California State Polytechnic University, Social Sciences Department, 1 Grande St, San Luis Obispo, CA 93407, USA.
E-mail: cmcarrig@calpoly.edu

consumption of the digital. These heuristics help us to draw comparisons between gender relations in data science and the manufacture of machine learning. Second, we explore what we call techniques of invisibility, a term we coined to describe patterns by which powerful actors in data science cultivate opacity to skirt oversight and conceal immoral and illegal acts (Burell, 2016; Pasquale, 2015). Techniques of invisibility refer to the ways in which an extreme imbalance between exposure and opacity, demarcated along fault lines of power, are fabricated and maintained, closing down the possibility for transparency in the production and applications of machine learning algorithms. Finally, we discuss how techniques of invisibility work to individuate workers and consumers, capture the resistance of social movements to exalt the digital economy as "revolutionary," and erase the contributions of its marginalized workers and the labor value of its users, exacerbating historical relations of inequality.

We ask: how does "consent" as an analytical framework help make visible the social dynamics and power relations in the capture, dissemination, and labor of data science knowledge? We argue that attention to gender in data science workplaces and human data extraction in broader social domains is a sound approach to rendering the social in computing visible and open to scrutiny and debate, with implications for public welfare and justice in the United States. While the social cannot be disaggregated from the technical, or practice from ideology, it may be fruitful to analyze consent in the particular social relations in this digital age where ideologies of the ruling class conflict with actual people's lived experiences and the conditions under which work to train machines to think and see is organized (Smith, 2004).

The impetus for this paper relates to its title, "The Revolution Won't Be Supervised," which is a quote from an esteemed data scientist exalting the promise of a popular trend in machine learning. Witnessing a white man appropriate Gil Scott-Heron's poem about the Black Power Movement while accepting a US$1m dollar award for technology that has negatively impacted Black communities in the United States (Benjamin, 2019) crystallized a pattern in our data that we wanted to understand. We were also curious why some in the field are calling for unsupervised machine learning and deregulation of data science outputs, while others, notably trans- and cisgender women,[1] and non-binary practitioners, are calling for increased oversight to harassment in the field and bravely discussing ethics and challenges of algorithmic outputs (Dickey, 2021; Hanna and Whittaker, 2020). These contradictory discourses related to oversight and accountability prompted us to explore a connection between harassment in data science workplaces and data extraction, which includes the capture and analyses of people's biological processes, behavior, communication, and social network patterns (Zuboff, 2019). Our comparative framework also unearthed more examples of the capture of the revolutionary spirit of social activism in the service of digital capitalism (Dean, 2009).

Consent and refusal

We frame consent not only as an individual "choice" but rather in a political context where "the society, not just the agent, is subject to critical scrutiny" (Sherwin, 1998: 37). In fact, consent frameworks in domains of reproductive politics often operate to authorize women's subjection and increase surveillance of reproductive decisions (Denbow, 2015). Operationalizing autonomy and consent from a feminist orientation can help us be mindful of concealed circumstances that structure our engagement with computing technology and the conditions in which this engagement occurs.

In the social domains in which Big Tech acquires the means to train machines to see and learn, do digital users have a choice to opt out of being surveilled? Privacy experts are concerned. Currently, a mechanism by which Big Tech companies gain consumers consent is a process of "notice and consent" governed by the 1980 statute Organization for Economic Cooperation and Development (OECD) Guidelines. OECD is considered too lax by some national and state-level lawmakers because it presents consumers with lengthy, complex notices regarding privacy protections (or lack thereof) and requesting they accept these terms or forego the use of the service offered (Cate and Mayer-Schönberger, 2013). It "makes opportunities to consent an unacceptable burden for most individuals" (Cate and Mayer-Schönberger, 2013: 68). For example, users in the United States encounter ~1462 privacy policies in a year, and each, on average, takes 10 min to read. It would take 76 working days to read through all of these agreements (McDonald and Cranor, 2008). Further, the expertise required to understand these agreements is uncommon and, too often, avenues to the resources required to gain this expertise are inaccessible to all but a privileged few (Margolis, 2008; Margolis and Fisher, 2002).

Consideration of privacy includes not just data collection, but also data use and dissemination (Nissenbaum, 2004). In fact, people's frustrations often stem from a sense of powerlessness, a feeling of "being forced" without knowing toward what ends it will be used (Andrejevic, 2014: 1682). The deluge of privacy agreements written in specialists' language that does not describe the dissemination terms of a user's data renders the consent recommended by OECD coercive. Like consensual sex is a form of collaborative activity, and not an object to be given (Kukla, 2020), consensual data relations would involve the technology providers treating consumers as "a partner in a mutually beneficial consensual relationship" (Wittkower, 2016: 13). Furthermore, notice and consent practices offer no option for negotiations. This is a social arrangement in

which one party acts unilaterally, and the other acknowledges and legitimates the vast inequalities between them by submitting to Big Tech's terms. These terms can include calcifying reductionist gender binaries, for example, in gender recognition software that nonconsensually assign a gender to users rather than recognizing self-professed gender-identities, terms that reproduce normativity and trans-exclusion (Keyes, 2018; Spiel et al., 2019). Zuboff (2019: 253) calls these unilateral arrangements a "dictatorship of no alternative." Big Tech offers two options: either submit to their terms or refuse technological products upon which the global economy depends (Sadowski, 2019).

Simpson (2007) describes indigenous conceptualizations of refusal as a means of maintaining sovereignty under colonial and neocolonial regimes. Benjamin (2016) augments this analytic frame, coining the term "informed refusal" to describe the possibility of people being able to opt out of biomedical research. In this article, we adopt informed refusal and further apply Simpson's scholarship to consider the conditions in which people have the right to refuse technology and technological intrusion (Benjamin, 2016). To understand one's opportunities for consent in a social relationship, we must "critically examine the stigma and penalties that may result from opting out" (Benjamin, 2016: 968). Big Tech's failures to provide alternatives to behavioral data extraction and create workplaces void of gender harassment incite a demand for alternatives. Refusal begets the possibility for doing things differently (Simpson, 2017). The costs of refusal can be borne collectively.

Our point here is not to make analogous two kinds of lived experiences (being harassed at work and being spied on via digital platforms) but rather to show a resonance resulting from workplace relations of dominance, which are then replicated in relations between digital merchants and their users. Both of these endemic problems—data consent and gendered dynamics in Big Tech—have been extensively studied, but in this paper, we do two novel things. First, we put the two in conversation looking for resonances and commonalities in the techniques used to normalize and perpetuate consent violations. Second, we do so from a feminist anthropological lens. Toward this end, we interrogate the social dimensions of machine learning and the powers they serve.

Feminist methodology

This paper is based on a subset of data from an ethnographic project called *Valuing the Social in Computing*, which investigates a hierarchy of knowledge within data science, comparing two subfields—human–computer interaction and machine learning—and their cultural practices and values regarding gender, race, and social applications. In the lineage of feminist anthropology, which reckons with the field's history of surveillance (Harrison, 1991), this article reflects on power and marginality and combats exploitation in support of organized efforts to transform "techbro culture" (Davis and Craven, 2016; Sharma, 2018: 4). A common trope in anthropology is to invoke the metaphor of the spyglass to describe the work we do (Hernández, 1996). On the one hand, the spyglass positions an anthropologist as an authoritative observer, rendering subjects vulnerable to objectification. On the other hand, the interpretative turn in anthropology, particularly from a feminist orientation, turned the spyglass around to interrogate the researcher and the powers she serves (Behar, 1995), and the powerful, e.g. "studying up" (Gusterson, 1997; Nader, 1972). Barabas et al. (2020) recommend that if data scientists also seek to lend their talent and skills to the social good, then powerful actors and the field's culture must come under scrutiny. Here we are studying up, using our spyglass to investigate surveillance and harassment—practices that elite members of data science actively conceal.

To analyze the gender relations in these sites of technoscience, and how human behavioral data extraction mirrors and reproduces aspects of these relations, we draw on a subset of data from our long-term project, including semi-structured interviews with those in the world of data science; public statements made by high-tech leaders; reports produced by Big Tech and their lobbyists; field notes taken at tech conferences; and the lived experience of the first author from her time as an insider in Big Tech. We did a domain analysis, analyzing the language of data scientists to understand relationships between their cultural concepts (Spradley, 1979). Themes related to "the social" emerged. We compared these themes with what marginalized members of machine learning said about the culture of their subfield in our interviews. There were striking contrasts between these sets of data that inspired our subsequent consideration of structural questions (Spradley, 1980) from a feminist orientation (Davis and Craven, 2016). As a feminist project, this article takes seriously the experiences of marginalized workers in Big Tech and unsettles the boundaries of insider/outsider, and self/other in research (Abu-Lughod, 1990; Davis and Craven, 2016).

Open secrets

To understand if there is a connection between workplace culture and the processes of machine learning, we are interested in what we call techniques of invisibility. Techniques of invisibility operate to help exempt Big Tech from ethical standards enforced in society, for example, protections for privacy and bodily autonomy. This contrasts with logics of workplace surveillance and data extraction of behavioral data, which demands unfettered exposure of users, a balance of scrutiny that Foucault (1995) argues favors those with excesses of privilege.

Techniques of invisibility do not erase powerful actors in data science from public view. Indeed, they are featured prominently on the world's stage. Instead, techniques of invisibility operate as ideological coordinates of value that desubjectivize these bourgeoisie actors, making secret their agency in the exploitation of workers and consumers. In other words, data science leaders use techniques of invisibility to displace attention from their modes of production and keep secret the sources of profit in the digital economy.

> "What covers secrecy?" … .Secrecy would thus be the practice in which the oppositions of private/public, inside/outside, subject/object are established, and the sanctity of their first term kept inviolate. And the phenomenon of the "open secret" does not, as one may think, bring about the collapse of these binarisms and their effects, but rather attests to their fantasmatic recovery
>
> (Miller, 1988: 207).

The fantasmatic recovery of the aforementioned binaries is an advanced form of the commodity fetish, in which algorithms materialize as "misrecognized social contracts" (Thomas, et al., 2018: 4). We argue that the fetish of algorithmic commodities leverage a hegemonic promise vested in them to forge and fortify the open secret of the source of their value. The two open secrets in data science that we are concerned with in this article are the gendered dynamics of power in its workplaces and one half of the dual process of Big Data—the extraction and circulation of behavior data, without which, the field of artificial intelligence would not exist. We argue that these open secrets operate in accordance with "black box" practices in computer programming, whereby, when an algorithm is too complicated to detail, engineers draw a box around it to signal its presence, and pay attention only to its inputs and outputs (Latour and Wollgar, 1986). In cultural practices related to algorithms, techniques of invisibility black box private interests and objectify both women data scientists and consumers of all genders alike, none of whom are given proper credit for being producers (producing inputs) of the "revolutionary" advances (upgrading outputs) in machine learning.

Mystification

How does data science come to be vested with hegemonic promise? In other words, how are the industry's aspirations transformed into decrees that divest billions of privacy rights? To make visible the assemblages required to craft "open secrets" in data science, it is important to pay attention to how data scientists frame the work they do and the impacts it has. "The selectivity with which system-builders represent their worlds … first [has] implications for the practitioners' concept of work for the way they go about knowledge acquisition; and second what this may imply in turn for the way expert systems function in the real world" (Forsythe and Hess, 2001: 29–30). There is a dichotomy between emic understandings of data science operations and public-facing discourse about data science products. Wolf (2019: 333) illuminates how machine learning developers view their everyday work with data as "hum-drum." However, the application of their work in the realm of the social has, according to a participant in this study, "almost mystical appeal … people are excited about what it can do!" So, on the one hand, the logic-based work of data science is mundane, and yet, on the other hand, its outputs are deemed capable of turning the world on its head. When data scientists frame the commodities of their labor objects of worship or religious veneration, it serves two purposes. One, it valorizes their expertise in the internal logics of an unsupervised machine learning algorithm, which are opaque due to the complexities of the methods used to create them (Burell, 2016). These specialized workers promote themselves as high priests of the digital era to appear exceptional, initiates who need an exclusionary sanctuary to create without questions or accountability. Two, the full scope of machine intelligence applications and its social implications remains shrouded in mystery, or even justified as natural. For example, the developers of deepfake technology acknowledge that perfecting their technology will push it past a threshold where either "something magical happens … or terrifying happens" (Fake Believe, 2019). If terror is a consequence of their magic, so be it; the developers expressed faith that tech will fix it (Fake Believe, 2019).

We argue that framing data science practitioners, their work, and outputs as mystifying is foundational to the power that computer elites wield in the US society. Techniques of invisibility are critical tools in data science's black box of magic tricks (Thomas et al., 2018) that make open secrets in data science possible. In the next section, we discuss two kinds of techniques of invisibility. The first we characterize as "epistemic injustice" (Fricker, 2007), which includes acts of ignoring, deletion, and obfuscation. The second, we call "the brotherhood," grouping together techniques regarding self-regulation, harassment, and worker surveillance. In each technique, we map the gendered dynamics in data science workplaces and the capture, extraction, and circulation of human behavior via the use of computing commodities. To do so, we borrow methodologically from Knorr Cetina (1999: 4), making the invisible visible by mobilizing social patterns in one domain of data science "as a sensor for identifying and mapping … patterns in the other."

Epistemic injustice

Epistemic injustice refers to structures of power that thwart epistemic practices of the marginalized and normalize this violence (Dotson, 2014; Fricker, 2007). We employ it

here to also include oppressive acts built into structures of power that do not prevent the subordinated subject from epistemic acts, but appropriate such acts and assimilate them into extant stratified hierarchies.

Techniques of invisibility

Acts of ignoring. Workplace culture in data science tolerates and black boxes the systemic problem of sexual harassment, which reproduces oppressive gender relations in the field (Carrigan, 2018). This may be a reason why data science remains stubbornly segregated, more so than more scientific fields (Cheryan et al., 2017). The permissiveness organizational leaders' show toward perpetrators when women report sexual harassment is the reason feminists often refer to this crime as an "open secret" in the US workplace (Cantalupo and Kidder, 2017). Much like the "prediction imperative makes individual ignorance the preferred condition for [data science's] rendition operations" (Zuboff, 2019: 253), so too does sexual harassment thrive when organizations' members and leaders practice "acts of ignoring" (Quinn, 2002). These acts of ignoring are predicated on assumptions (rooted in institutional relations of gender and race) of who is a competent knowledge producer and who is not, whose bound of privacy and autonomy are respected and whose are transgressed. These assumptions are generated in a social matrix of domination and disempowerment along vectors of race, gender, and sexuality. For example, women of color scientists are targeted at higher rates than white female peers (Clancy et al., 2017) and gendered racial stereotypes shape how "women in technical positions manage their gender and sexuality" (Alfrey and Twine, 2017: 30–31).

An example of how data scientists practice acts of ignoring was documented by data scientists (Lum, 2017). Lum courageously details her experiences at a tech conference with a predatory male professor, whom she refers to as "S." "S" harassed and attempted to assault her and regularly made degrading, sexually explicit comments about junior women in his field to the delight of his male colleagues. "To say that S's bad behavior is an open secret in the community is an understatement" (Lum, 2017). Years later, several professors witnessed "S" sexually assault another woman colleague. One witness responded by saying that he would have to find a way to caution his female students on how to "how not to get raped by S" so that he does not lose any gifted students (Lum, 2017). Rather than censure his predatory peer, this male scientist instead put the onus of responsibility on individual students, maintaining the culture of silence around harassment and assault. His silence is a form of epistemic injustice maintaining the "open secret" of sexual assault and allows predators to continue hurting women and driving them out of data science (Lum, 2017). Lum refused to persist in academic data science and credits her attrition to multiple instances of harassment and its tolerance. Harassment not only impacts individual targets' lives, it makes women in data science more vulnerable. Scant representation increases visibility and heightens scrutiny, creating conditions ripe for surveillance. This power dynamic in regard to representation correlates with boundary violations in data extraction practices. The people most negatively impacted by decisions around surveillance and privacy are the least likely to have a seat at the table where these decisions are designed (Broussard, 2018; Noble, 2018).

Deletion. To inquire into another dimension of epistemic injustice, we ask: where does Big Data come from? To begin to answer this, we need to tease out the two dimensions of Big Data—data extraction and data analysis. Forsythe and Hess's (2001) ethnography of data science describes the extraction process for training machines to learn. Forsythe's work brings our attention to aspirational dimensions of machine learning extolled in technoscientific culture and the "deletion" of certain labors, namely the social, intellectual, and material labor outside the direct action of problem-solving (Forsythe and Hess, 2001). By deletion, we mean a pattern of selectivity that does not consider critical elements of a system, such as, say, sexism, or surveillance.

At the time of Forsythe's research (more than 20 years ago), machine learning required acquiring knowledge from human "experts" to transfer it into machines. This created a bottleneck of expert knowledge. Current trends in machine learning attempt to solve this bottleneck, training machines to make deductions not from expert knowledge alone, but from oodles of human data recorded via digital devices. In other words, today, the labor needed to train machines to reason is no longer just coming from "experts" in data science (who actively choose to participate in this data collection), but from anyone who consumes computer technology and engages with social media platforms, i.e. the majority of the US society. A data scientist in academia put it this way: "In our current technological age we are no longer just content consumers, we are content providers" (McMullen and Waisome, 2019). The source of this geyser of raw materials is often deleted from discourses on machine intelligence.

Currently, the amount of data mined by Big Tech is staggering, with 1.7 MB of data being generated every second for each person on the earth (IBM, 2017). For consumers of platforms such as Facebook, the data collected are neither benign nor scant. It is personal. Curran (2018) discovered that Facebook had collected roughly 600 MB worth of data from him, including contacts, all files, images, and messages he had ever sent, his location from every place he had logged on and the information from each of the applications he had ever linked to his account (Curran, 2018).

Why is it important to collect vast swaths of personal details from people? Yann LeCun, director of Artificial Intelligence (AI) research at Facebook and one of the three white men awarded the 2018 Turing Award, discussed in his keynote at the 2019 the Annual ACM Symposium on Theory of Computing (STOC) conference the limitations of machine learning. He argued that the "salvation" to overcoming these constraints is "self-supervised learning … training very large neural networks to understand the world through prediction. Self-supervised learning is going to be revolutionary and the revolution will not be supervised" (first author field note, June 2019).[2] Prediction accuracy depends on the amount of data fed to the machines. LeCun makes no mention of where this data comes from but, given his place of work, it would not be unreasonable to assume this data comes from consumers of his company with billions of users such as Dylan Curran. LeCun is not an outlier in his failure to acknowledge the data collection process for training sets. Even socially conscious conversations in the field about machine learning processes typically begin after knowledge is extracted from users. Yet, machine learning would not exist without this data whose sources are deleted, much like women who pioneered computing have been erased from its celebrated history (Abbate, 2012; D'Ignazio and Klein, 2020; Ensmenger, 2010; Hicks, 2017).

Big Tech leaders are still silencing women in their companies. Satya Nadella, Microsoft chief executive officer (CEO), told a room full of thousands of women at the 2014 Grace Hopper Celebration of Women in Computing to not ask for raises and, instead, wait for "karma to reward them" (Larson, 2014). This advice illustrates another component of the deletion form of techniques of invisibility. It is not enough for dominant group members to devalue worker's labor value. The workers themselves are expected to delete their own value from the equation of Big Tech's success and the bounty of its rewards. The economic context of women's work as data scientists is also deleted here, as Nadella takes "flight into the misty realm of religion" (Marx, 1976/1990: 165), where, according to Nadella, women are to "have faith that the system will give you the right raise" (Larson, 2014). Note how both LeCun and Nadella invoke religious discourse when discussing machine learning, which could help explain why data scientists find this area of study "mystical." The value digital users and minoritized workers contribute to Big Tech's power and profits is thus deleted and a fetishizing fable of salvation is placed in its stead.

Obfuscation. Deletion not only propagates epistemic injustice, but also clears a space upon which Big Tech leaders can project their prevarications. We call this technique "obfuscation," which also devalues the sources of knowledge and value required to train machines. Obfuscation differs from a deletion in that it "involves deliberate attempts at concealment" (Pasquale, 2015: 6). For example, Zuboff (2019: 172–173) coined the term "Surveillance as a Service" to describe the veneer of benevolence shellacking how the human behavioral data are collected by digital devices, "including texts, emails, GPS coordinates, social media posts, Facebook profiles, retail transactions, and communication patterns."[3] Much like paternalistic expressions of care favored by colonial rulers (Murphy, 2015), "Surveillance as a Service" speaks to how Big Tech frames its extraction processes as a benefit.

For example, at the 2019 Turing Award ceremony at STOC, the moderator extolled the virtues of machine learning, claiming that "billions of people benefit from computer vision and speech recognition, which create new tools for astronomy, medicine, robotics and material sciences" (first author field note, June 2019). However, the social harm it causes, and will cause, widely discussed in public and scholarly discourses (Benjamin, 2019; Nafus, 2018; Noble, 2018) was elided. Thus, not only are the surveillance mechanisms used to train machines deleted from public discourse, but also are their negative consequences obfuscated, e.g. ad-targeting algorithms that perpetuate racial and gender discrimination in housing and employment (Ali et al., 2019).

Some of Big Tech's leaders help us peek behind the curtain of data acquisition for machine learning. For example, Satya Nadella, Microsoft CEO, stated: "The opportunity we have in this new world is to find a way of catalyzing this data exhaust from ubiquitous computing and converting it into fuel for ambient intelligence" (Zuboff, 2019: 163). "Data exhaust" is a euphemism for behavior data. It is a metaphor, similar to "data mining" that misrepresents data practices in a way that belies agencies, intimacies, and significance (Kerssens, 2019). Further, it defines this source of profit as waste, which makes sense only using imperial logic by which the colonizer "liberates" the peasant from the soil, and makes the land productive, land that, if left to indigenous people's stewardship, would fail to be "productive" and therefore wasteful (Pateman and Mills, 2007). Obfuscation is, thus, predicated on settler colonial logic of expropriation.

Other worrisome examples of obfuscation are commonplace, the likes of which we need look no further than Facebook. In her keynote address at the 2011 Grace Hopper Celebration of Women in Computing, Sheryl Sandberg, the Chief Operating Officer of Facebook, said:

> If you think about what changes our lives, the great things that have changed lives over the course of history, there's political movements and there's technology, and technology is what's driving now … You think about this social revolution where each one of us gets technology to power who we are as individuals, I don't think there is anything which has more impact
>
> (Sandberg, 2011).

That computerized technology galvanizes social revolutions separate from politics is not only an amusing claim

from a powerful executive who hosts heads of state in her workplace (Sandberg, 2013: 27), it is a strategic obfuscation of the influence her corporation has on politics around the world. Since this keynote, the public has now glimpsed inside the black box of Facebook's modus operandi. The company she helms collaborates with "personality mercenaries" such as Cambridge Analytica and world governments to reorient the capacities of machine learning systems from futures markets toward "guaranteed outcomes in the political sphere" (Zuboff, 2019: 281).

Behavioral data extraction is opaque by design and the "revolutionary" mythos surrounding machine intelligence is key to its obfuscation. Sandberg's claim, "technology is driving now,"[4] is a form of myth-making that removes agency from the production of data science and accountability for its outputs (Benjamin, 2019). By obfuscating labor and the decision-making power that humans have in algorithmic technology, the public is denied the opportunity to see how values, views, ideologies, assumptions, and desires do, in fact, animate sociotechnical infrastructures. Note too how Sandberg also uses a vehicle metaphor reminiscent of Scott-Heron's (1971) sentiment in "The Revolution Will Not Be Televised," in which he declares: "the revolution will put you in the driver's seat." This evinces Big Tech's efforts to capture the revolutionary spirit of US social movements in service of its own capitalist ends. Akin to her colleague LeCun's revolutionary rhetoric, who also cribs from Scott-Heron's poem, Sandberg's framing process reinforces the unassailable power and moral status of Big Tech and obfuscates the role that people play in the cognitive training of machines. Those driving the digital economy are unregulated and untaxed, creating machines built from surveillance practices that defy scrutiny.

Sandberg adopted another moral stance in her book *Lean In* (2013), promoting a neoliberal feminist agenda that tells women to "internalize the revolution" (Rottenberg, 2018; Sandberg, 2013: 11). This message acts as a form of epistemic injustice, which delegitimizes ways to address structural inequalities in women's lives and makes us responsible for our own oppression (Rottenberg, 2018). This epistemic violence protects harassers, and stymies change toward egalitarianism in Big Tech. The individuating rhetoric of neoliberal feminism is another instance of the co-optation of social movements by Big Tech leaders who wish to control how resistance is performed—if it is at all. Obfuscation, as a technique of invisibility, is key to amplifying the belief that data science serves the common good, reproducing it with authority pillaged from collective actions for justice such as feminism, or the Black Power movement.

The Brotherhood

Another dimension of techniques of invisibility relates to the fabrication of data science as an elite, all-male cabal, a Brotherhood as it were. Marked by secrecy and exclusivity, it not only blocks access to non-dominant group members, but also weeds out behaviors and learning styles that do not conform to masculinized logic and "rigorous" rationality (Riley, 2014; Turkle and Papert, 1990). Blocking and weeding maintain elitism, which begets prestige, which works in tandem with gendered ideologies to externalize reproductive labor to people minoritized in—to paraphrase bell hooks (2000)—white supremacist and capitalist heteropatriarchy.[5] Google's firing of Timnit Gebru, a Black computer scientist, and her white co-lead Margaret Mitchell from its AI Ethics team is a prime example of how Big Tech weeds out those who do not submit to white, heteropatriarchal power relations in its workplaces and dare to question its outputs as anything less than beneficent (Hanna and Whittaker, 2020; Metz, 2021).

Set apart from the profanity of materiality of life and those assigned to care for this aspect of being human, the Brotherhood, in its sanctity, can set its own rules, maintain privacy from the prying eyes of subordinated folks, and coopt revolutionary progress made by social movements such as feminism and the Black Power movement in the United States to credential its corporate activism.

Techniques of invisibility

Self-regulation. To understand the context in which data scientists proclaim unsupervised machines are revolutionary, we must first consider debates on regulation of science. That scientific endeavors are separate from society remains a popular belief (Franklin, 1995; Jasanoff, 2011; Nader, 1972, 1996; Winner, 1986). For some scientists, "understanding the science" and "understanding how science works" serve as threshold tests for the right to question the socio-ethical dimensions of technology (Jasanoff, 2011: 634). Too often, the scientific community has responded negatively to regulation and justifies opacity in its knowledge production with claims of public illiteracy of scientific specialization. Further, technology companies have lobbied across the United States to deregulate biometric data laws and privacy protection (Zuboff, 2019). For example, the Information Technology Industry (ITI) Council (2017), a lobbying group for Amazon, Facebook, and Google, warned lawmakers to "keep their hands off their algorithms" (De Vynck and Brody, 2017). The ITI Council demanded a "flexible regulatory approach," telling governments that they should "use caution before adopting new laws, regulations or taxes that may inadvertently or unnecessarily impede the responsible development and use of AI" (ITI 2017: 1). While vaguely acknowledging the liabilities of AI technology, Big Tech's lobbyists also stressed that the companies' algorithms should remain proprietary. Governments must not require access to corporate algorithms operating practices for AI should be self-regulated (De Vynck and Brody, 2017).

Deletion, obfuscation, and self-regulation operate as techniques of invisibility that maintain and reproduce the asymmetries of power between Big Tech and their customers, individuate responsibility for consent and refusal in regard to citizens' data privacy and designate the social not as a site of commonweal but rather a site of secrecy and extraction.

Homosocial bonding and harassment. Self-regulation advocates frame algorithms as wholly benevolent, whose creators need not be held accountable. The corollary to this in the domain of gender in data science is the bystander who treats harassment with an insouciance born from patriarchal authority. Harassment operates to call attention to women's presence in the data science workforce and makes their status exceptional. Simultaneously, it makes data science a homosocial space. Unlike religious sects such as the Catholic Church, the "Brotherhood" in data science does not explicitly bar women from participating but instead, consolidates male power through hegemonic labor value and practices (Carrigan, 2018).

Some men regard women as interlopers who "brazenly" dare to tread in this homosocial space. For example, at my former high-tech job, I, the first author, was meeting with the vice president of IT, an ex-Marine. He directed me to go to his office and use his phone to call my boss, another vice president, to get his buy-in on a high-priority issue. The phone had caller identification and Dick, also a former military officer, picked up the phone and greeted me: "Hello, Captain Boss Man—how are ya, Jughead?" When I introduced myself, he yelled at me: "What are you doing using Joe's phone?" He resented that I had caught a glimpse of male camaraderie and interrupted the bonding that he took for granted in brotherhoods such as the military and Big Tech. He refused to engage with me on the consensus question I called about and, soon after, I was denied a promotion to a leadership position. This incivility and retribution signaled I did not belong.

This experience heightened my theoretical sensitivities as to why men harass women in data science. A man who is hostile or creepy toward his female colleague wants to remind her of her subordinated status in broader society. This status can be characterized by patriarchal norms demanding that women exist to service men, particularly in the realms of sexuality and reproductive labor (Kukla, 2020; Manne, 2017). In technoscience, women's talents are deleted and diminished; they are not peers, they are interlopers deserving of conquest. For example, one research participant said: "When I came to grad school, I was a target. I had advances, full-on sexual advances by three guys in the first month." Further, harassment in data science not only signals to women that they do not belong, it punishes other transgressions, including competing with men for "masculine-coded prizes" and robbing them of privacy to bond in homosocial settings (Manne, 2017: 117). Harassment hurts individuals and collectively disenfranchises women from the digital economy.

Gender harassment is a near-daily experience for some women. Male peers insult them with comments such as: "Who did you sleep with to get this internship?" or "You know you're a diversity hire, right?" One researcher, who has been in this field a decade, finds "it infuriating to be asked whether I am a recruiter, or a 'plus one,' or whether I did the work myself" (Else, 2018: 611). In these ways, men's harassment of women colleagues in data science mirrors the asymmetries of power between users and tech corporations in the collection of behavior data, whereby the dominant group maintains power through enforcing insider/outsider status and objectifying the "outsider" group as incompetent with little to no claim to privacy and boundaries.

Expropriation is also noticeable in data science workplaces. For example, in our study, research participants have reported men stealing their ideas, taking credit for their work or giving credit to other men for women's work, a status-seeking act of prostration used to fortify men's homosocial bond at women's expense. Sexism happens at organization levels too, resulting in lost promotion opportunities. A research participant described her mentor's experience in "organization full of these young, hotshot male software engineers … they ignore what she says all the time, and she always gets overlooked for the hotshot male managers. Back when she was at [a Big Tech company], she got skipped over for a couple of promotions and that's why she quit." Even though this data scientist had been in the field for over 30 years, she was denied access to leadership and wealth accrual in favor of men with less experience. The speaker's repeated use of "hotshot" speaks to the men's status in the organization and implicitly, her mentor's exclusion from this sanctified brotherhood, even though she had far longer tenure at the corporation and more experience than they did.

Women in these situations fear retaliation against those who report, a common employment condition that increases the prevalence of gender harassment (National Academies, 2018). This calls into question data scientists' ability to refuse and resist harassment when their livelihood and careers are at stake. Retaliation against individuals who report, acts of ignoring, and homosocial bonding among men are some techniques of invisibility that allow gender harassment to remain entrenched in data science (National Academies, 2018). These techniques of invisibility in data science do not merely hide incivility and illegal acts, they also encourage community members to perform ignorance of exploitative relations. The consequent open secrets generate tailwinds for cisgender, white men to ascend to positions of prominence in data science. These techniques mirror behavior data collection in that they are a farcical version of consent, "which may be better termed acquiescence" (Sadowski, 2019: 7). Like end-user agreements,

harms incurred by harassment are, too often, the individual target's responsibility to bear, a culture phenomenon that resists negotiation for amelioration, instead, offers a choice to either acquiesce or leave the field.

Worker surveillance. For digital workers', rights to privacy are nullified when they sign hiring documents (Pasquale, 2015). Employee sales, correspondence, behaviors, and schedules are analyzed for efficiency, productivity, and potential malfeasances (Pasquale, 2015; Van Oort, 2019). However, workers have neither access to the data collected on them, nor the opportunity to gaze back at those collecting data on them. Further, computing merchants and the complex computational infrastructure they control, remain unsupervised due, in large part, to aggressive lobbying efforts. In this way, tech leaders are like priests in patriarchal religions such as, say Catholicism, whereby, through the tradition of confession, they know the secrets of their devotees but remain mysterious and unaccountable (Daly, 1985).

The increasing ability of corporations, businesses, and governments to unilaterally peer into the lives of individuals is a stark example of the one-way spyglass of surveillance that leaves those being surveilled susceptible to control. When considered through a refractive surveillance framework, in which the effects of surveillance extend to populations beyond the ostensible target (Levy and Barocas, 2018), the nullification of a worker's right to privacy may have implications for consumers too.

Workplace surveillance also may also support the conditions that permit the open secret of gender harassment. Anonymity is critical in protecting those who come forward with experiences of harassment from retaliation, ostracization, and other forms of backlash, especially when corporations have had their data hacked (Dredge, 2015). However, the surveillance of workers' emails and interactions with others means that instances of harassment may already be recorded. If an employee comes forward, those who analyze the data extracted from workers may already be aware of the incident and may dissuade targets from reporting.

Women in data science are aware of being monitored, both in the workplace and beyond. Participants in this study took great care with the formal documentation of their lived experience as data scientists. Some would not allow us to use their workplace email for fear that any exchange that was less than glowing about their employer would trigger retaliation. Others refused to use digital platforms connected to the "cloud" to, again, protect against retribution. These acts of care and fear speak to the dangers of critiquing sexism in computing domains and the dangers of the surveillance capacities of technoscience. Our research participants, creators of these technoscientific infrastructures undergirding our society, understand the capacities of surveillance best—and its dangers.

Institutional regulation

The lack of options for workers to resist harassment sans retaliation helps us better understand whether or not digital consumers have a meaningful choice to refuse being surveilled by Big Tech. For this to be possible, regulations would need to protect workers and consumers from boundary transgressions regarding privacy, autonomy, and respect. Supervising Big Tech must be institutionalized, addressing the cultural context in which individuals make decisions to engage in it, whether as a minoritized member of data science or as a consumer of digital devices and platforms. Scientific institutions must put measures in place that allow people actually to refuse without recrimination and not leave up to the individual to question and challenge powerful institutions (Benjamin, 2016). For example, Institutionalized practices such as regulatory algorithm audits could mitigate bias and enforce accountability (Engler, 2021). Corporations balk at regulatory oversight and assure us that they have their own systems in place for regulation and society's best interests at heart, but these systems can fail and internal reviews operate much like techniques of invisibility, directing attention away from powerful actors' culpability.

Recently, there have also been efforts to scrutinize sexism in data science and regulate data extraction processes used to generate Big Data. For example, scientists in theoretical computing organized a collective response to harassment, creating standards for reporting harassment and oversight to curb retaliation in their field (Irani et al., 2018). In regard to data privacy, in 2018, the European Union enacted a robust piece of regulation titled General Data Protection Regulation, requiring the transparent, fair, and lawful collection of data (*Regulation*, European Union). It demanded that the use of this data be limited to specified purposes and the storage of such data limited to only the necessary amount of time needed for the completion of said purposes (*Regulation*, European Union). In 2020, California enacted similar privacy laws called the California Consumer Privacy Act. These regulatory measures are laudable in their efforts to protect privacy and data rights. However, they pose complex, technical difficulties for compliance, which diminish their ability to be legally enforced (Goldwasser and Park, 2017).

Discussion

Our comparative optics on knowledge politics in data science extends the claim that algorithms manifest cultural values and meanings (Seaver, 2017) to include sexism in high tech and its ripple effects on society. The recent mantra in data science—the revolution will not be supervised—has meaning beyond technical domains. It is an ethos that pervades the social dimensions of data science as well, with implications for relations of power that produce algorithms

and their broader ambits of influence. This ethos is the latest example of how "brogrammers [have] sought to monopolize this power through an alpha male affect in the workplace that encourage[s] the technological ends and aims of an 'elite' subgroup" (Hicks, 2013: 87). Past research has revealed how sexist norms and values are imported into computing classrooms, labs, and workplaces (Carrigan, 2018; Cheryan et al., 2009). Here we have inverted this methodological approach to understand how sexist culture in data science is exported to society and to what effects. In this article, we contribute a feminist perspective to how the algorithm is conceptualized and circulated to enforce broader rationalities about how society should function and to whose benefit (Beer, 2017).

Techniques of invisibility operate to aggrandize some aspects of machine learning and mask others in efforts to control the terms of the debate over technology. Algorithmic governance nudges workers in data science and digital consumers to acquiesce, indeed submit, to boundary transgressions. We have described whose lives and behaviors are made visible and made vulnerable to this kind of coercion and how dominant groups make their behaviors and institutional practices opaque and thus, nearly unassailable. The absence of bidirectionality of the spyglass of transparency is a critical aspect to relations of power in the information age.

Further, the lack of institutional oversight and regulatory policies gives immunity to computing corporations, their algorithmic assemblages and their employees who sexually predate their colleagues for any harm they cause. This places an undue burden on individuals to decide whether or not a company's policies sufficiently protects their privacy, or whether or not to risk retaliation to report the harassing behavior of a colleague or superior. A culture of open secrecy regarding harassment in data science workplaces creates a culture of open secrecy that tolerates and even encourages boundary transgressions in data collection and circulation.

Finally, the surveillance of digital users, who are at once the consumers and workers who provide content for machine learning, and the concomitant deletion of their labor value, is reminiscent of settler colonial practices of governance and exacerbates exploitative labor relations and disparities of access, wealth, opportunities, and resources in the global economy. Feminist surveillance studies and critical algorithmic research are helping us better understand how "the use of surveillance practices and technologies [work] to normalize and maintain whiteness, able-bodiedness, capitalism, and heterosexuality, practices integral to the foundation of the modern state" (Dubrofsky and Magnet, 2015: 7).

Conclusion

Using feminist methodologies and frameworks of consent, we have compared the open secret of how data science acquires the means by which to train machines to the open secret of gender harassment in the field. Machine learning requires certain actors' labor go unrecognized and, like harassment, occurs under clandestine conditions that may violate consent, autonomy, and reasonable expectations of privacy. We described techniques of invisibility, the practices by which powerful actors in data science refuse accountability, enabling their behaviors and outputs to remain unsupervised and unaccountable. These techniques may be useful in making further connections between epistemic violence, sexism and surveillance, sussing out persistent boundary violations in data science, debunking Big Tech's mythologies and combatting its harmful practices.

Further, the harassment of women in data science, the appropriation of revolutionary rhetoric, and the obfuscation of human agency in machine learning signify a concerning lack of care for the social in technoscience—e.g. social relations, social impacts, and social accountability. This dereliction and subsequent boundary violations are "so grotesque that it is impossible to see how anyone who really thinks about it lives with it—and yet, we do" (Bridle, 2019: 2). Yet, consider these issues we must, and we can begin by measuring meaningful consent by the capacity for informed refusal. "An informed refusal … is seeded with a vision of what can and should be, not only a critique of what is" (Benjamin, 2016: 970). Right now, compliance in regard to gender relations and human behavior data collection is secured coercively, without enthusiastic consent. Our vision is to redesign ubiquitous computing whereby codes of conduct and consent are transparent and data science institutions and its leaders are supervised by public mechanisms of accountability.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by an National Science Foundation Engineering Education and Centers Grant NSF 1751314.

ORCID iDs

Coleen Carrigan https://orcid.org/0000-0002-2930-891X
Madison W Green https://orcid.org/0000-0003-4353-681X

Notes

1. Our view of "women" is not dichotomous, but rather, meant to include members of socioeconomic classes that have been marginalized on the basis of gender and sexuality.
2. LeCun's attributed "The Revolution Will Not Be Supervised" slogan to Dr. Alexei Efros. Efros gave a 2017 talk at the

International Conference on Computer Vision called "The Revolution Will Not Be Supervised." Efros' title slide had a picture of two people from the African diaspora making the Black Power fist and breaking out of chains.

3. We find merit in the critiques of Zuboff's *The Age of Surveillance Capitalism*, particularly those that challenge its naturalization of capitalism and lack of grounding in Marxist scholarship and surveillance studies (Ball 2019; Morozov 2019). However, Zuboff's thorough explication of the process of data capture and extraction from digital users is, in the context of this study, useful. It demystifies Big Data, framing it instead as a manufacturing process, and lays bare the extent of datafication and its stakes.
4. Please see Sandberg's quote above.
5. Sharma (2018) describes this externalization of reproductive labor in Big Tech as "working from Mommy's basement," a co-optation of the politics of care in feminist movements.

References

Abbate J (2012) *Recoding Gender: Women's Changing Participation in Computing. History of Computing*. Cambridge: MIT Press.

Abu-Lughod L (1990) Can there be a feminist ethnography? *Journal of Feminist Theory* 5(1): 7–27.

Adam A (1998) *Artificial Knowing: Gender and the Thinking Machine*. Florence: Routledge.

Alfrey L and Twine FW (2017) Gender-fluid geek girls: Negotiating inequality regimes in the tech. *Gender and Society* 31(1): 28–50.

Ali M, Sapiezynski P, Bogen M et al. (2019) Discrimination through optimization: How Facebook's Ad Delivery can lead to biased outcomes. In: *Proceedings of ACM on human-computer interaction (CSCW)*, November 2019, vol. 3, Article 199, pp. 1–30. New York, NY: ACM.

Andrejevic M (2014) The Big Data divide. *International Journal of Communication* 8: 1673–1689.

Ball K (2019) Review of Zuboff's the age of surveillance capitalism. *Surveillance & Society* 17(1/2): 252–256.

Barabas C, Doyle C, Rubinovitz JB et al. (2020) Studying up: Reorienting the study of algorithms around issues of power. In: ACM conference on fairness, accountability and transparency, Barcelona, Spain, 27–30 January 2020.

Beer D (2017) The social power of algorithms. *Information, Communication & Society* 20(1): 1–13.

Behar R (1995) Introduction: Out of exile. In: Behar R and Gordon D (eds) *Women Writing Culture*. Berkeley: University of California Press, pp. 1–29.

Benjamin R (2016) Informed refusal: Toward a justice-based bioethics. *Science, Technology, & Human Values* 41(6): 967–990.

Benjamin R (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge, MA: Polity.

Bridle J (2019) The age of surveillance capitalism by Shoshana Zuboff review—we are the pawns. *The Guardian* (accessed 1 July 2020).

Broussard M (2018) *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge, MA: MIT Press.

Burell J (2016) How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society* 3(1): 1–12

Cantalupo NC and Kidder WC (2017) Mapping the title IX Iceberg: Sexual harassment (mostly) in graduate school by college faculty. *Journal of Legal Education* 66(4): 850–881.

Carrigan C (2018) 'Different isn't free': Gender @ work in a digital world. *Ethnography* 19(3): 336–359.

Cate FH and Mayer-Schönberger V (2013a) Notice and consent in a world of Big Data. *International Data Privacy Law* 3(2): 67–73.

Cheryan S, Plaut VC, Davies PG et al. (2009) Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology* 97(6): 1045–1060.

Cheryan S, Ziegler SA, Montoya AK et al. (2017) Why are some STEM fields more gender balanced than others? *Psychological Bulletin* 143(1): 1–35.

Clancy KBH, Lee KMN, Rodgers EM et al. (2017) Double jeopardy in astronomy and planetary science: Women of color face greater risks of gendered and racial harassment. *Journal of Geophysical Research: Planets* 122(7): 1610–1623

Curran D (2018) Are you ready? Here's all the data Facebook and Google have on you. *The Guardian*.

Daly M (1985) *The Church and the Second Sex*, 3rd edn. Boston: Beacon Press.

Davis DA and Craven C (2016) *Feminist Ethnography: Thinking Through Methodologies, Challenges, and Possibilities*. London, UK: Rowman & Littlefield.

De Vynck G and Brody B (2017) Amazon, Google Lobbyists Warn Regulators to Keep Their Hands Off AI. *Bloomberg*, 18 November, 20.

Dean J (2009) *Democracy and Other Neoliberal Fantasies: Communicative Capitalism and Left Politics*. Durham, NC: Duke University Press.

Denbow J (2015) *Governed Through Choice: Autonomy, Technology and Reproductive Politics*. New York: NYU Press.

Dickey MR (2021) Google fires top AI ethics researcher. *Techcrunch*. https://techcrunch.com/2021/02/19/google-fires-top-ai-ethics-researcher-margaret-mitchell/ (accessed 21 February 2021).

D'Ignazio C and Klein LF (2020) *Data Feminism*. Boston: MIT Press.

Dotson K (2014) Conceptualizing epistemic oppression. *Social Epistemology* 28(2): 115–138.

Dredge S (2015) Why the workplace of 2016 could echo Orwell's 1984. *The Guardian*.

Dubrofsky RE and Magnet SA (2015) *Critical Interventions: Feminist Surveillance Studies*. Durham, NC: Duke University Press.

Else H (2018) Can conference shed its reputation for hosting sexist behaviour?: AI meeting wants to become more inclusive, but survey suggests it has a long way to go. *Nature* 563: 610–611.

Engler A (2021) Auditing employment algorithms for discrimination. Brookings Institute, Center for Technology Innovation.

Ensmenger N (2010) *The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise. History of Computing*. Cambridge: MIT Press.

Fake Believe (2019, November 22). *The New York Times*. https://www.nytimes.com/2019/11/22/the-weekly/deepfake-joe-rogan.html (accessed 7 September 2021).

Forsythe D and Hess DJ (2001) *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence*. Stanford, CA: Stanford University Press.

Foucault M (1995) *Discipline and Punish: The Birth of the Prison*. New York: Vintage Books.
Franklin S (1995) Science as culture, cultures of science. *Annual Review of Anthropology* 24: 163–184.
Fricker M (2007) *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
Goldwasser S and Park S (2017) *Proceedings of the 2017 on Workshop on Privacy in the Electronic Society*, Dallas, TX, 30 October 2017, pp. 99–110. New York, NY: Association for Computing Machinery.
Gusterson H (1997) Studying up revisited. *PoLAR: Political and Legal Anthropology Review* 20(1): 114–119.
Hanna A and Whittaker M (2020) Timnit Gebru's exit from Google exposes a crisis in AI. *Wired*. https://www.wired.com/story/timnit-gebru-exit-google-exposes-crisis-in-ai/ (accessed 10 January 2021)
Harrison F (1991) Ethnography as politics. In: Harrison F (ed) *Decolonizing Anthropology: An Anthropology for Liberation*. Arlington: American Anthropological Association, pp. 88–112.
Hernández G (1996) Multiple subjectivities and strategic positionality: Zora Neale Hurston's experimental ethnographies. In: Behar R and Gordon D A (eds) *Women Writing Culture*. Berkeley: University of California Press, pp. 148–165.
Hicks M (2013) De-brogramming the history of computing. *IEEE Annual History of Computing* 35(1).
Hicks M (2017) *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing*, 1st edn. Boston: MIT Press.
Hooks B (2000) *Feminism is for Everybody: Passionate Politics*. Boston, MA: South End Press.
IBM (2017) Turning Big Data into action. IBM Watson Advertising.
Information Technology Industry Council (ITI) (2017) *ITI AI Policy Principles Executive Summary*. 1–6. https://www.itic.org/resources/AI-Policy-Principles-FullReport2.pdf
Irani S, Chambers E, Blum A et al. (2018) Report from the Ad hoc committee to Combat Sexual Harassment and Discrimination on the Theory of Computing Community: Theory of Computing.
Jasanoff S (2011) Constitutional movements in governing science and technology. *Science and Engineering Ethics* 17: 621–638.
Kerssens N (2019) De-agentializing data practices: The shifting power of metaphor in 1990s discourses on data mining. *Journal of Cultural Analytics*.
Keyes O (2018) The misgendering of machines: Trans/HCI implications of automatic gender recognition. *ACM Human Computer Interaction* 2, *CSCW*, 22.
Knorr Cetina K (1999) *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.
Kukla QR (2020) A nonideal theory of sexual consent. *Ethics* 131(2): 270–292.
Larson S (2014) Microsoft CEO Satya Nadella to women: Don't ask for a raise, trust karma. October 9, PARS International Corp. Available at: https://www.businessinsider.com/microsoft-ceo-satya-nadella-to-women-dont-ask-for-a-raise-trust-karma-2014-10 (accessed 15 January 2015)
Latour B and Woolgar S (1986) *Laboratory Life: The Construction of Scientific Facts*. Princeton, NJ: Princeton University Press.
Levy K and Barocas S (2018) Refractive surveillance: Monitoring customers to manage workers. *International Journal of Communication* 12: 1166–1188.
Lum K (2017) Statistics, we have a problem. *Medium*.
Manne K (2017) *Down Girl: The Logic of Misogyny*. Oxford, UK: Oxford University Press.
Margolis J (2008) *Stuck in the Shallow End: Education, Race, and Computing*. Cambridge, MA: The MIT Press.
Margolis J and Fisher A (2002) *Unlocking the Clubhouse: Women in Computing*. Cambridge, MA: The MIT Press.
Marx K (1976/1990 [1867]) *Capital*, Vol. I. London: Penguin Classics.
McDonald AM and Cranor LF (2008) The cost of reading privacy policies. *Journal of Law and Policy for the Information Society* 4(3): 543–568.
McMullen K and Waisome J (2019) Making history every day: Privacy, Security, and Authorship In *Modern figures podcast*, *ep. 009*. Available at: http://modernfigurespodcast.com/making-history-every-day-episode-009/
Metz C (2021) Who is making sure the A.I. machines aren't racist? *New York Times*.
Miller DA (1988) *The Novel and The Police*. Berkeley: University of California Press.
Morozov E (2019) Capitalism's new clothes. *The Baffler*. 4 February 2019.
Murphy M (2015) Unsettling care: Troubling transnational itineraries of care in feminist health practices. *Social Studies of Science* 45(5): 717–737.
Nader L (1972) Up the anthropologist: Perspectives gained from studying up. In: Hymes DH (ed.) *Reinventing Anthropology*. New York: Pantheon Books, pp. 284–311.
Nader L (ed) (1996) *Naked Science: Anthropological Inquiry in Boundaries, Power and Knowledge*. New York: Routledge.
Nafus D (2018) Exploration or algorithm? The undone science before the algorithms. *Cultural Anthropology* 33(3), pp. 368–374.
National Academies of Science, Engineering, and Medicine (NASEM) (2018) *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine*. Washington, DC: The National Academies Press.
Nissenbaum H (2004) Privacy and contextual integrity. *Washington Law Review* 79.
Noble SU (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: NYU Press.
Pateman C and Mills CW (2007) *Contract and Domination*. Cambridge: Polity Press.
Pasquale F (2015) *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
Quinn BA (2002) Sexual harassment and masculinity: The power and meaning of "girl watching". *Gender and Society* 16(3): 386–402.
Regulation (EU) 2016/679.
"The Revolution Will Not be Supervised" Promises Facebook's Yan LeCun (2018) *Tandon*. Available at: https://engineering.nyu.edu/news/revolution-will-not-be-supervised-promises-facebooks-yann-lecun-kickoff-ai-seminar.
Riley DM (2014) What's wrong with evidence? Epistemological roots and pedagogical implications of "evidence-based practice" in STEM education. In: *2014 ASEE Annual Conference & Exposition, Indianapolis*, IN, 15-18 June 2014,

pp. 24.1373.1–24.1373.9. Washington, DC: American Society of Engineering Education.

Rottenberg C (2018) Women who work: The limits of the neoliberal feminist paradigm. *Gender, Work & Organization (Special Issue)* 26(8): 1–10.

Sadowski J (2019) When data is capital: Datafication, accumulation, and extraction. *Big Data and Society* 6(1): 1–12.

Sandberg S (2011). *Keynote Address. Grace Hopper Celebration of Women in Computing*. Portland: Anita Borg Institute.

Sandberg S (2013) *Lean in: Women, Work, and the Will to Lead*. New York: Alfred A Knopf.

Scott-Heron G (1971) The revolution will not be televised. *Pieces of a Man*.

Seaver N (2017) Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4(1): 1–12.

Sharma S (2018) Going to work in mommy's basement. Once and future feminist. *Boston Review*, 19 June.

Sherwin S (1998) Relational approach to autonomy in health care. In: Sherwin S (ed) *The Politics of Women's Health: Exploring Agency and Autonomy*. Philadelphia: Temple University Press, pp. 1–19.

Simpson A (2007) On ethnographic refusal: Indigeneity, 'voice,' and colonial citizenship. *Junctures: The Journal of Thematic Dialogue* 9: 67–80.

Simpson A (2017) The ruse of consent and the anatomy of 'refusal': Cases from indigenous North America and Australia. *Post Colonial Studies* 20(1): 18–33.

Smith DE (2004) Ideology, science and social relations: A reinterpretation of Marx's epistemology. *European Journal of Social Theory* 7(4): 445–462.

Spiel K, Walker AM, DeVito MA et al. (2019) Queer(ing) HCI: Moving forward in theory and practice. In: CHI conference on human factors in computing systems, Glasgow, UK, 4-9 May 2019. Glasgow, UK: Association for Computing Machinery.

Spradley JP (1979) *The Ethnographic Interview*. Ann Arbor, MI: Holt, Rinehart and Winston.

Spradley JP (1980) *Participant Observation*. Ann Arbor, MI: Holt, Rinehart and Winston, p. 195.

Thomas S, Nafus D and Sherman J (2018) Algorithms as fetish: Faith and possibility in algorithmic work. *Big Data and Society* 5(1): 1–11.

Turkle S and Papert S (1990) Epistemological pluralism: Styles and voices within the computer culture. *A Journal of Women in Culture and Society* 16(1): 128–157.

Van Oort M (2019) The emotional labor of surveillance: Digital control in fast fashion retail. *Critical Sociology* 45(7): 1167–1179.

Winner L (1986) *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: The University of Chicago Press, p. 214.

Wittkower DE (2016) Lurkers, creepers, and virtuous interactivity: From property rights to consent to care as a conceptual basis for privacy concerns and information ethics. *First Monday* 21.

Wolf CT (2019) Conceptualizing care in the everyday work practices of machine learning developers. In: Designing interactive systems (DSI), San Diego.

Zuboff S (2019) *The Age of Surveillance Capitalism: The Fight for A Human Future at the New Frontier of Power*. New York: Public Affairs.