Race and digital discrimination

Dr. Seeta Peña Gangadharan, Associate Professor, Department of Media and Communications, London School of Economics and Political Science,
s.p.gangadharan@lse.ac.uk

Summary:

Race and digital discrimination is a topic of interdisciplinary interest that examines the communicative, cultural, and social dimensions of digital technologies in relationship to race, racial identity, and racial inequalities, harms, or violence. Intellectual traditions in this area span vast terrain, including those that theorize identity and digitally mediated representation, those that explore social, political, and economic implications of unequal to technological resources, and those that explore technical underpinnings of racial misidentification in digital systems. The object of inquiry thus varies from racialized interactions in digital spaces, to the nature or extent of access to high-speed broadband infrastructure, to levels of accuracy in computer vision systems. Some research orients towards policy or technical interventions to safeguard civil and human rights of individuals and groups and prevent racial discrimination in the design and use of digital technologies. Other strands of race and digital discrimination scholarship focus on diagnosing the (both recent and distant) past to excavate ways in which race itself functions as a technology.

The variety in approaches to the study of race and digital discrimination has evolved organically. Following a general concern for bias in the design, development, and use of digital technologies, scholarship in the 1990s began to center its attention on the problem of racialized discrimination in computerized, data-driven systems. In the earlier part of the 1990s, scholars writing about surveillance warned about the social, political, and economic consequences of sorting or categorizing individuals into groups. Towards the latter half of the 1990s, several scholars began scrutinizing the incorporation of specific values—and hence

bias—into the computational design of technological systems, while others began looking explicitly at racialized interactions among users in virtual community and other online space. Throughout the early 2000s, scholarship—particularly in European and U.S. American contexts—race and racialization in different aspects of design, development, and use of digital technologies began to emerge. The advancement and rapid commercialization of new digital technologies—from platforms to AI—has heightened interested in race and digital discrimination alongside social movements and social upheaval in relation to problems of systemic and institutionalized racism. Scholars have also taken interest in examining the ways in which race itself functions as a technology, primarily with attention to race's discursive power.

The study of race and digital discrimination in all its varieties will remain relevant to issues of social ordering and hierarchy. Scholarship on race and digital discrimination has been instrumental in broadening critical and cultural perspectives on technology. Its ability to expose historically and culturally specific dimensions of race and racial inequality in digital society has helped scholars question modernist assumptions of progress and universal benefit of technological development. This body of work will continue to push discussion and debate on the nature of racialized inequalities in future eras of technological innovation.

Keywords

algorithmic bias, algorithmic racism, cyberhate, data and discrimination, data-driven discrimination, decolonial AI, digital discrimination, digital redlining, electronic redlining, fairness in machine learning

Introduction

The study of race and digital discrimination spans an impressive range of interrogations of the interplay between race, racialized populations, and digital technologies. While a significant portion of this corpus focuses on the social construction of race and technology's role in those processes, a sizable set of works adopt a technology-centered approach, exploring constructionist elements of digital technology and the digitization or datafication of social life through the lens of race, racial difference, and racism. Across the past several decades, scholars have maintained or renewed interest in racialized differences in access to digital technology and representation in digitally mediated spaces, often though not exclusively at the intersection of issues of governance, law, and ethics.

Overview

The study of race and digital discrimination skews heavily towards U.S., North American, or European contexts. With its unique history of structural racism and a long tradition of broaching race in the study of media and communication technologies, the United States served as an early origination point for scholarship on race and digital discrimination. In the 1980s throughout the early 2000s, a great deal of this work focused on access, expression, identity, and the internet. Several key works also examined targeting, predation, surveillance, and database or data-driven systems. Literature on barriers to entry of minority populations to science, technology, engineering, and mathematics emerged in the 1990s, though with variable levels of scrutiny of the issue of discrimination. In the second decade of the 2000s, and following U.S. and global uprisings for racial justice and racial equity, race and digital discrimination literature has bloomed, with particular attention to algorithmic racism or algorithmic oppression. Though the evolution of the U.S. Civil Rights movement and complementary racial equity and racial justice movements has inflected scholarship

concerned with law, ethics, and governance, race and digital discrimination studies are equally moored in histories of colonialism, postcolonialism, and decoloniality.

On the European continent, the study of race and technology has its own trajectory. European powers played unique, strategic roles in the development of imperialism, colonization, and discourses on race. As well, postwar development of a fundamental rights frameworks in the wake of Nazism and the Holocaust led to European nations eliminating the use of race in politics. Race neutral policies stand in stark contrast to the U.S. American context where racial categorization is accepted and informs civil rights discourse on the need for identifying race to rectify racial inequities, otherwise known as affirmative discrimination. Europe also has its own specific history of enslavement, persecution, and genocide of the Romani population, a fact which locates the discourse on race (and technology) beyond considerations of empire.

These particularities factor into the scholarship on European race and digital discrimination. Early European scholarship on race, internet access, and digital inclusion tends to mirror interests in found in North America literature. Access, expression, and racial identity received—and continues to receive—noteworthy attention in the European context, and the discriminatory impacts of uneven broadband deployment, unfair broadband pricing models. However, European scholars charted a unique path of study of race and automated technologies, informed in part by concerns for race-neutrality or "discrimination awareness." Beginning in the late 1990s, European computer scientists broached the topic of discrimination when contemplating the impacts of automated computer systems. By the second decade of the 2000s, this literature has evolved to a focus on bias in machine learning or bias in AI (Pedreschi et al., 2008). Meanwhile, European scholarship on data justice and data colonialism tend not to theorize race, though intersect and implicate conversations on

race and digital discrimination (Couldry & Mejias, 2019; Marjanovic et al., 2021; L. Taylor, 2017).

Outside of the North American and European contexts, scholarship on race and digital discrimination thrives, reflecting distinctive histories of race and racism, as well as technology adoption (see also, Banaji & Bhat, 2021; Matamoros-Fernández & Farkas, 2021; Udupa et al., 2021). The regional varieties of race and digital discrimination scholarship may not achieve the same visibility as prominent U.S. authors on race and digital discrimination (see, for example, Benjamin, 2019; Brock Jr., 2020; Noble, 2018), but they importantly reveal the unique interplay between history, social norms and values, and legal and technological developments.

Having outlined regional varieties, we can now turn to specific domains of inquiry. Networked technologies (such as World Wide Web, internet, social media, platforms) and automated technologies (e.g., intelligent or smart technologies, machine learning systems, predictive analytics, data-based or data-driven systems) dominate discussion of race and digital discrimination. As alluded to above, different areas of scholarship anchor their work in theories of race or racial difference, discrimination and violence, and digital technology to varying degrees.

Commented [A1]: AU, from OUP: Whenever possible, please rephrase here and throughout the manuscript to avoid time-sensitive language such as "now" or "today." Refer instead to specific years or decades, or "early 21st century."

Commented [A2R1]: I'm a little confused with this comment, as the "now" refers to the ordering of arguments in the chapter, as opposed to temporal ordering and the topic being examined. Thus, I have not changed this text.

Race as technology, racial technologies, carceral technologies

Race functions as a system of power or a way to organize economic, social, and political power. The idea that race is a technology (or, race-as-technology) extends this line of thinking. As Chun (2009) argued, the race-as-technology rubric centers attention on how individuals and groups are racialized, rather than what constitutes race. The rubric serves as a conceptual guide, emphasizing that race is examined as a technology, as both tool and symbol. It invites us to employ the same framework developed by historians and philosophers

of technology to unpack context and contingency in technological development and use. That is, just as humans and technology coevolving, culture and biology also coevolve, with deep implications for how society "does" race. "[N]ot only can we theoretically and historically better understand the force of race and technology and their relation to racism[,] we can also better respond to contemporary changes in the relationships between human and machine, human and animal, mediation and embodiment, nature and culture, visibility and invisibility, privacy and publicity" (Chun, 2009, p. n.p.). Chun argues that by framing race as technology, the question of *how* individuals and groups are racialized takes precedence over the question of *what* is race. Rather than debating the boundary between biology and culture, the study of race as technology is preoccupied with the processes and practices which determine such boundaries as well as their responses.

Numerous studies have emphasized the co-productive relationship between race and technology. Using the lens of visual culture and, more specifically, artistic practices, Gonzalez (2006) scrutinized white-as-default standards in digital imaging and image production technologies. Nakamura (2007) identified the interface of a webpage as the space where digital bodies are categorized, gendered, and racialized. Poster (2019) examined the racializing logics of platforms and their users, arguing that a racialized, othering gaze of users emerges in the foreground and background of consumption and exchange of digital services. the co-construction of race and platform services. Nopper (2019) compared between credit scoring and social media-based lending tools, excavating the narratives of discrimination and antidiscrimination circulating between them. Scannell (2019) argued predictive policing technologies enable or make racialization and criminalization possible. The point here is that the carceral imaginary and digitally networked or platform-based technologies arise together.

Studies using the race-as-technology rubric have doubled down on explorations of the "how" and, on some occasions, begun to ponder "what if." To the latter, Coleman (2009)

interpreted race-as-technology framework with respect to moral philosophy. Her argument centers on agency and the positioning of racialized self in relation to or in spite of repressive systems of power. Drawing from Kant, McLuhan, and others, Coleman stressed mutability or the fact that race is socially constructed as technology, and liberation. Namely, as a social construction, race-as-technology can be envisioned as insurrectionist, rather than as inherited and destined and can serve as a mechanism to open new possibilities of freedom, autonomy, and self-representation (see also, Coleman, 2004). The emphasis on agentic processes has subsequently informed other explorations of race, gender, and the body (Gill-Peterson, 2014; Keeling, 2014), visual culture (Raengo, 2013), and agency, racialization, and technoculture (Hochman, 2014; Johnson, 2016).

To the former, Noble's (2018) analysis deconstructs search algorithms, augmenting her analysis with Black feminism and political economic theories. Specifically, she highlighted how the insensitivities of Google engineers combined with the company's profit-minded motives to perpetuate racism, sexism, exploitation, and extremism among information seekers online. The biased, bottom-line politics of search engines stands in stark contrast to company marketing, which pitches search technology as a leveller for all individuals and groups in society. Benjamin (2019) looked at the ways in which technologies are created within a racist context and deepen existing racial inequities in the United States. New digital tools, such as predictive policing, facial recognition, and other biometric systems, extend the state's practice of mass incarceration of Black populations in the United States, especially Black men. Like pre-digital carceral tools, technologies of the 21st century other, alienate, and ensnare Black bodies. Like Noble, Benjamin critiques neutrality or objectivity in popular discourse about technological invention and introduces the term carceral technologies to pinpoint the deleterious racializing impacts of modern-day tools.

Race, technology, and affirmative discrimination

In terms of usage or access of digital technology, including internet-based technologies, the problem of unfair treatment against individuals and groups based on racial categories has provoked a vibrant discourse on affirmative discrimination in teaching and learning about technology. An affirmative approach to digital discrimination posits that typically disadvantaged groups ought to be targeted with educational interventions to improve digital literacy or, more broadly, knowledge of science, technology, engineering, and mathematics (STEM). Doing so allows disadvantaged groups to rise above systemic challenges that make them less equal in accessing and using technology. As compared with the constructionist race-as-technology approach mentioned above, the affirmative approach tends to be rooted in behavioral frameworks and often is justified in relation to national educational strategies and global competitiveness. For example, in the United States, the push for improved STEM education ties to developments in agriculture, aeronautics and space, and ICTs and the maintenance of world-leading reputations (H. B. Gonzalez & Kuenzi, 2012).

Within the field of media and communication, studies of digital skills, inclusion, literacy, and equalities/inequalities often overlap with STEM, though with slightly different emphasis. Seminal studies in the field often responded to facile calls for increasing internet access without heeding institutional and cultural conditions that facilitate internet use and benefits. The explicit affirmative framing was often more muted than in STEM scholarship. For example, works by Hargittai (2002), Warschauer (2002), van Dijk and Hacker (2003), Livingstone (2004), Selwyn (2004), and Buckingham (2006) established the importance of examining digital literacy as a component of inclusion and meaningful participation in digital society. Subsequent studies extended analysis to deepen awareness of sociodemographics, psychodemographics, and historical and cultural contexts that impede or enhance positive outcomes due to technology use (Livingstone & Helsper, 2007; van Deursen et al., 2016).

However, by and large seminal studies agreed to the positive benefits of internet use and the importance of remedying disparities in access and use, especially those experienced by already marginalized groups.

Specific technologies have come under the purview of digital literacy studies, with racial disparity or its remedy being an occasional topic of central concern. More often than not, problems of racialized marginalization and racial discrimination are used as motivation for study of user knowledge or behavior and explorations of literacy efforts. Klawitter and Hargittai (2018), Zarouali, Helberger, and Vreese (2021), and Shin, Rasul, and Fotiadis (2021) cite discrimination and bias as justification for the investigation into algorithmic skills and knowledge (see also, Hobbs, 2020).

Since champions of affirmative discrimination first articulated STEM principles, scrutiny of the underlying assumptions of STEM and digital literacy initiatives, in particular their capacity to address racial disparities, has grown. Bullock (2017) claims that STEM programs in the United States fit into a longer, troubled history of racialized urban renewal plans that often lead to displacement and dispossession of Black families and people of color. Upon evaluation STEM policies Bullock found that White, middle-class families target STEM schools and drive gentrification to the disadvantage of residents of color (see also, Morales-Doyle & Gutstein, 2019). Crooks (2022) questioned the extent to which digital inclusion efforts have, by and large, failed to buoy racialised populations in the United States. Looking at young people outside the United States, Banaji (2015) wrote about negative effects of technocentric optimism found in Western digital media discourse has damaging spillover effects, arguing it overhypes technological solutions to historically specific problems faced by youth in the Global South.

Nevertheless, scholarship on STEM and digital literacy remains fertile ground to explore affirmative discriminatory approaches to technology education. As internet

shutdowns proliferate throughout the world, pandemic lockdowns continue or reappear, and concerns about disinformation interact with political, social, and economic change, and the costs are borne by society's most marginalized groups, affirmative approaches will remain a vital object of inquiry.

Race, technology, and negative discrimination

While affirmative discrimination involves purposefully discriminating against certain groups for their betterment, negative discrimination refers to the disparate harmful or hurtful treatment of groups based on racialized differences. Most conspicuously, research on negative discrimination, race, and digital technology began with broadband infrastructural concerns and tracks with the appearance of other new technologies, from platforms to machine learning systems, and more.

Digital redlining, the "digital divide," and digital inclusion

Beginning in the mid 1990s, the term electronic redlining was used to describe the exclusion of Black people, indigenous people, and people of color (BIPOC) in the United States from the benefits of information access and communication in telephony, telecommunications, and internet markets. Civil society groups petitioned policymakers to prevent redlining in the telecommunications market and popular discourse embraced the idea (Howard, 1995; Randolph, 1995). Kahl (1997) examined civil society voices highlighting racism in telecommunications markets that allow firms to bypass minority neighborhoods. Plascencia (1999) examined electronic redlining in "video dial tone" markets. Baynes (2004) crafted a legal analysis market barriers faced by Native Americans, Latinos, and African Americans in the 1990s and early 2000s, including legacy regulations that hindered these populations from accessing basic telephony services. Research highlighting racial disparity and telecommunications markets has continued to the present day, ranging from digital

subscriber line (DSP) (2008) to fiber (2017), and from rural to municipal broadband markets (Dailey et al., 2010; Galperin et al., 2021). Focusing on intentionally discriminatory technology policies and practices, Nadaf (2021) stressed the paralyzing effects of government internet shutdowns, which cut off Kashmiri students and teachers from learning and teaching opportunities

Since an initial focus on telecommunications markets, research has looked beyond broadband network infrastructure, to apps, services, or content, as well as political participation in the governance of digital systems. Some of this research builds upon debates of the digital divide or digital inclusion. For example, Gangadharan (2012) looked to earlier examples of redlining in the early 20th century—or low-tech data profiling—examples to warn of impacts on racialized populations. Taylor and Sadowski (2015) noted the ways in which lenders use Facebook data to determine creditworthiness. Admonishing the effects of surveillance capitalism and platformization in the United States, Gilliard (2017) employed the term digital redlining to refer to structure, content, and the politics in between, explaining that redlining refers to a process "of doing difference, a 'doing' whose consequences reinforce class structures" (n.p.). Friedline, Naraharisetti, and Weaver (2020) and Friedline and Chen (2021) wrote about the impacts of poor rural broadband infrastructure on poor people and communities of color hoping to gain access to fintech tools in the United States.

Some of the literature strays from a digital inclusion focus, charting its own terms of reference for discussion of discrimination and redlining. Noble's work, mentioned above, employs the term "technological redlining" with reference to racist and sexist search algorithms. In general, Noble (2018) contends that sorting mechanisms present in software and on platforms have the power to exclude members of society, especially those from marginalized groups. Noble's work relates to a class of scholarship equally concerned with

hate, violence, and aggression directed towards Black people, people of color, ethnic minorities, migrants, and indigenous groups.

Discriminatory technologies, computer bias, and discrimination-aware data mining

Other research on digital redlining looks beyond digital inclusion to issues of discrimination, privacy, and surveillance, three areas that are distinct, yet overlap. The earliest works in this vein targeted 20th century innovations in marketing technologies segment and discriminate between populations. In ground-breaking research, Gandy (1993) used the term "discriminatory technologies" to denote the power of classification wrought by U.S. companies operating vast electronic databases. His work brought nuanced understanding of how private corporations—and not just the state—develop and benefit from surveillance complex. Lyon (1994) also unpacked the subtleties of surveillance and their technologies in society, moving further away from dominant narratives of Big Brother. Meanwhile, Regan (1995) looked at electronic databases and other technologies, with focus on the idea of privacy as a public good.

By the next decade, surveillance studies, critical geography, legal studies, communication, and other fields pondered the consequences—and often deleterious and dehumanizing impacts—of databases and data-driven technologies. Of particular interest were technologies that determine individual value or worth and that individualize risk in various contexts. Analyses spanned social welfare (Gilliom, 2001; Monahan, 2008), policing (Harcourt, 2007; Schauer, 2003), security (Amoore, 2008), and health (Roberts, 2009). Continuing his emphasis on the political economic aspects of discriminatory technologies, Gandy (2009) covered a vast array of market and non-market environments, including health, education, housing, and financial services, warning that new data mining techniques in computerized databases exacerbate already existing social, racial, and economic divisions in society.

While U.S. legal scholars and social scientists were probing databases, privacy, surveillance, and discrimination, computer scientists also began inquiring into the discriminatory impacts of automated decision making. As early as the 1990s, the topic of bias appeared in scholarly investigations into computer automated systems (Friedman et al., 1996; Friedman & Nissenbaum, 1993, 1996). In the U.S., computer scientists primarily concerned themselves with computer ethics and values, and not race or antidiscrimination per se (Friedman, 1996, 1997). European computer scientists eventually started to explore the social issues already raised by legal scholars and social scientists: namely, the intersection of computational techniques and privacy and discrimination. These researchers worried that individuals who were otherwise anonymous in databases could be reidentified or de-anonymized and differentially treated in the process of discovering patterns in datasets (Pedreschi et al., 2008). What resulted from this worry was an influential set of studies on discrimination-preventing or discrimination-aware data mining and knowledge discovery in databases (Berendt & Preibusch, 2012; Calders & Verwer, 2010; Corda, 2007; Kamiran & Calders, 2009; Pedreschi et al., 2009, 2008).

From algorithmic discrimination to algorithmic fairness

In the 2010s, U.S. computer science, legal studies, and social science converged in their interest in data, computer automated systems, and social impacts, with concerns focusing on algorithms and algorithmic power (boyd & Crawford, 2012; Edelman & Luca, 2014; Mulligan & Dwork, 2013; Sandvig et al., 2014; Sweeney, 2013). The push to research algorithmic practices, processes, and impacts came on the heels of a report on big data, privacy, and innovation published by the Obama Administration (U.S. Executive Office of the President, 2014). The report cautioned that high-tech discrimination could result from automated systems that targeted people based on patterns discovered in massive data sets, threatening to supersede privacy challenges and usher in a new era of redlining in the digital

economy. To support this claim, the White House report relied on an array of legal, science, advocacy research, and investigative journalism, including an in-depth examination of transparency in automated decision systems (Pasquale, 2015), a report of microtargeting and consumer scoring (Dixon & Gellman, 2014), and testimonies warning against profiling of vulnerable communities (Crawford & Schultz, 2014; Croll, 2012; U.S. Executive Office of the President, 2014).

Not too long after, research under the rubric of automated or algorithmic discrimination blossomed. Key among these literatures is Selbst and Barocas' study (2016) highlighting the ways in which computer automated systems—now referred to as automated decision systems—could lead to disparate impacts on subgroups or vulnerable populations. The article knitted technical literatures on data mining with legal analysis of antidiscrimination implications and unpacked the ways in which algorithmic decision making may reflect unintentional prejudice towards groups. Other analyses of algorithmic bias or algorithmic discrimination ranged from a focus on programmable accountable measures (Kroll et al., 2015) to institutionalized racism (Noble, 2018), from lack of minority representation in industry (Garcia, 2016) to algorithmic redlining (Lambright, 2019).

An equally powerful research rubric focused on fairness in automated or data-driven systems flourished in the 2010s. Like their European counterparts, computer scientists in the United States worried about privacy and discrimination. However, U.S. computer scientists offered a rubric of fairness, as opposed to discrimination. Key works were committed to discovering mathematical models to emulate "fair affirmative action" and treating different groups similarly (Dwork et al., 2012; Feldman et al., 2015; Friedler et al., 2016; Venkatasubramanian, 2019) and to protect privacy through fairness interventions (Dwork, 2017). Over time, however, the development and application of facial recognition led some "fairness" scholars to highlight problems of misidentification of gender and racial features.

Buolamwini and Gebru (Buolamwini & Gebru, 2018) argued that high rate of misclassification of faces of Black women as compared with those of White men meant that facial recognition systems would deny racialized populations the benefits they deserved, if not harm them (see also, Raji et al., 2020).

While the emerging field of computational notions of fairness has been met with great enthusiasm in computer science, interdisciplinary publications and interdisciplinary citational practices point to a growing concern that technical interventions may require historical and contextual considerations in the design and deployment of so-called fair automated systems (Selbst et al., 2019) (Selbst et al., 2019). Computer scientists and social scientists continue to challenge algorithmic fairness or computational fairness (Blodgett et al., 2020; Gangadharan & Niklas, 2019; Hanna et al., 2020; Hoffmann, 2019; Hutchinson & Mitchell, 2019; Sambasivan et al., 2021; Sánchez-Monedero et al., 2020; Stevens & Keyes, 2021).

Identity, representation, discourse

Works interested in representation, expression, censorship, and harm comprise a sizable subfield on race and digital discrimination. Common foci include identity formation and representation of racialized populations in digitally mediated environments, the technical or structural features shaping the experience of racial identity or racial difference, or racist speech and discourse in digitally mediated environments.

While the focus on identity, representation, and discourse shares commonality with the race-as-technology rubric, the two areas differ by virtue of a focus on *what* rather than *how*. Scholarship centered on identify, representation, and discourse often considers what technology does to racial identity, racial difference, and racism, as well as what racism, racial difference, and racial identity do to technology. By contrast, the race-as-technology rubric is interested in how society does race or how people and institutions racialize and other through or with technology.

Race and deconstructing the myth of digital democracy, internet equality, digital postracialism

Though early internet studies tended to shy from examinations of race, there are notable exceptions which deconstruct race in cyberspace against the backdrop of techno-utopian discourse and highlight racial stereotyping, violence, or harms. These studies approach digital discrimination in its social and cultural forms.

Looking at the construction of Asianness in multi-user domains, Nakamura (1995) used the term identity tourism to describe the ways in which privileged users adopt a racial identity different from their own and perform racially stereotyped behavior to appear authentic. The narrow range of stereotyped behaviors conveys the idea of the internet as post-racial, where free choice reigns supreme and no one is confined by structural forces or racial realities. Nakamura's later work uses visual cultural theories to unpack racial construction and deconstruction of digital identity as well as sense-making and interpretation of racialized identities online (Nakamura, 2007).

Critique of the internet-as-equalizer discourse continues to thrive, with scholarship featuring a wide range of interest in the violence and hate thriving in networks and platforms. Daniels' (2008) initial foray into White supremacy online decried claims of colorblindness in internet culture by systematically detailing the digitally mediated spaces, practices, and impacts of hate speech. The myth of colorblindness extends to gaming spaces, where whiteness is routinely privileged and racial stereotypes are reinforced in gaming culture (Condis, 2021; Daniels, 2015; Daniels et al., 2012; Gray, 2012; Nakamura, 2017). Studies of digital media and post-racialism, the idea that race and racism have become irrelevant and outdated or that society has and should retreat from race, criticize technology companies for exaggerating claims of diversity and inclusion (Noble & Roberts, 2019) and using proxies for racial identification online (Powell, 2018).

Histories of digital technology

Technological development is always political (Winner, 1986), and racist norms and practices feature routinely in the history of digital technologies' politics. In the early days of the development of personal computing and digitization more generally, engineers designing operating systems embraced the language of master and slave when labelling disk operating systems (Sinclair, 2004). As noted by Chun (2019), the fields underlying modern-day data science developed from racist rubrics of social Darwinism and eugenics. Writing from a decolonial perspective, several computer scientists and cognitive scientists have also linked current-day innovations to white supremacist and imperialist frameworks of the 19th and 20th centuries (Cave & Dihal, 2020; Katz, 2020; Mohamed et al., 2020).

By contrast, there is also a body of work that pushes back against the discriminatory focus of much literature on marginality and technology and centers the role of marginalized people in histories of invention and innovation. In the United States, research on this topic emerged partly as a response to "digital divide" and its overfocus on historically disadvantaged users. Uplifting Black, Latino, indigenous, and other minority scientists and technologists, research in this vein details the hidden accomplishments of Black and Brown women in science and technology fields across the United States (Taborn, 2008). The rubric of design justice approaches inventiveness and innovation from an affirmative position, advocating for the involvement of members of marginalized communities in the co-creation of technological artefacts and infrastructures (Costanza-Chock, 2020). Design justice centers the experiences and needs of groups sitting at the intersection of multiple injustice.

Black culture and cultural power also figure prominently in works that depart from the idea of needy, disadvantaged users. Networked and other digital environments thrive as a result of the cultural norms, values, and practices that racialized groups bring to online spaces. A body of work focused on research on race, racial identity, and the role of culture in

shaping networked and other digital environments rethinks innovation histories. Steele's (2021) exploration of Black digital feminism focuses on Black women's creativity in repurposing otherwise oppressive technologies, including innovations in self-presentation online. Brock's (2020) notion of distributed Blackness emphasizes the intentional and agentic aspects of identity formation thanks to the affordances of digital and online spaces. While Steele, Brock, and others acknowledge the conditions which constrain Black self-expression, they stress the inventive ways in which a culture of Blackness survives and thrives despite power imbalances.

Hate speech and violence

Hateful, racist speech acts can result in discriminatory harms, such as the exclusion of certain groups from gainful employment, denial of access to services, or stigmatization and cultural alienation. A key concern is that hate speech encourages hate crimes (see Berlet, 2001), and debates on hate speech focus on the legal permissiveness of hate speech acts, often pitting free expression against other social norms and values, such as civility, respect, and recognition (Dworkin, 1977; Kenyon & Scott, 2020). Early investigation into online hate speech argued that new online speech practices and new affordances of online spaces put pressure on lawmakers to reconsider of traditional legal measures designed to protect hate speech (Leets, 2002). Later research has shown the widening impacts of online hate speech on emotional, mental, and physical well-being as well as the chilling effects that such speech induces in the targets or victims of hate speech (Gelber, 2021; Gelber & McNamara, 2016; see also, Keipi et al., 2017). The study of online hate and discriminatory harms thus ties to questions about digital governance, including the impracticalities of state monitoring of hate speech, which results in increased attention to social responses and civic tools (Citron & Norton, 2011; Iginio et al., 2015), detection and prevention by machine learning systems

(Burnap & Williams, 2015; Fortuna & Nunes, 2018; Schmidt & Wiegand, 2019), and platform responsibility (Dijck et al., 2018; Helberger et al., 2018).

As with other domains of study of race and digital discrimination, research on online hate speech is frequently evaluated against the promise of digital democracy as well as claims that digital technologies are neutral. The "platformization" of speech puts platform companies at the center of decision-making with wide-ranging social, economic, and political consequences (Gillespie, 2010). The growth of new digital technologies has brought with it many undemocratic practices (Morozov, 2011), which are instantaneous, borderless, and often anonymous. Others have demonstrated the role of platform companies in distributing and amplifying racist, user-generated content and the racist impacts of vague strategies of platform governance (Ben-David & Fernández, 2016; Matamoros-Fernández, 2017).

Meanwhile, scholars from across the globe are broadening a subfield that historically focuses on North American—and in particular U.S.—contexts of hate speech (Matamoros-Fernández & Farkas, 2021). The effort to de-Westernize studies of extreme speech brings the study of digitally mediated hate speech beyond the confines of liberal democracy or the democratic promise of the internet. As argued by Pohjonen and Udupa (2017; see also, Udupa et al., 2021), historical particularities and cultural specificities shape the way that hate speech—or extreme speech, as conceptualized by the authors—develops and circulates online. Starting from local context allows nuanced understanding of the interactions between internet cultures, speech cultures, and regulatory environments (Pohjonen & Udupa, 2017). Discourses on civility, for example, have colonial roots, and this colonial legacy continues to inform how extreme speech is practiced in everyday contexts online (Udupa et al., 2021). Unique legal histories of antidiscrimination, antiracism, and migration also shape the interplay between hate speech, racialization, and discrimination (Hervik, 2019). In addition to socio-political contexts, users' values, online behavior, and interactions with contemporary

contexts warrants close investigation, not only to discover impacts on oppressed and minoritized communities that bear the brunt of violence and discrimination wrought by online hate but also to assess and improve solutions designed to counter online hate (Banaji & Bhat, 2021).

Racial capitalism, race, and global capitalism

There is rising interest in the analysis the specific forms of racialized information or surveillance capitalism, as well as the intersections between race, global capitalism, and digital society. A focus on the forms of capital accumulation in relation to systemic racism and racial violence brings to light the less-than democratic impacts of digital technologies.

To this end, Chakravartty and her collaborators have attempted to situate the extractive and expropriative forms of global, digital capitalism alongside histories of racial violence and imperialism. In an investigation into predatory inclusion in global financial systems, Chakravartty and Silva (2012) unpacked the demonization of racialized groups as unfit economic subjects. While their analysis treads lightly on the role of digital technologies, their scrutiny of predation and responsibilization paves the way for later analyses which further elaborate racist violence and imperialism in the context of so-called digital freedoms (Aouragh & Chakravartty, 2016) and media and information infrastructures (Chakravartty, 2019).

Taking a slightly different tack, other scholars lend more attention to private actors who dominate the technology industry and who influence the behavior of states. Building off of Pasquale's (2015) black box technologies, McMillan Cottom's (2020) use of the term obfuscation helps elucidate the ability of technology companies to not only obscure how their products and services work and but also defend such technologies, extractive and exploitative as they may be. Furthermore, obfuscation interacts with and intensifies predatory inclusion that drives digital, racial capitalism (McMillan Cottom, 2020). Kwet's (2019) examination of

digital colonialism locates hegemony of Western technology companies and imperial state surveillance in a legacy of racist colonial practices. His perspective on digital colonialism thus differs slightly from the concept of data colonialism, which refers to the deep asymmetries between data owners and data subjects, on the one hand (Thatcher et al., 2016), and datafication as a new stage of colonialism, on the other (Couldry & Mejias, 2019). Writing in the tradition of media and film theory, Beller's (2021) examination of computational racial logics posits that computation is racialized value extraction, easily detected in realms such as policing, credit, and housing.

Commented [A3]: AU, from OUP: Please discuss, briefly, the main threads in scholarship on your topic (i.e. include a Discussion of the Literature section), including past approaches to the subject as well as research questions that remain or that are currently being pursued. While this cannot serve as an exhaustive historiography, it should discuss, in broad strokes, the scholarship on your topic.

Please also include a selected bibliography of the 5-10 sources (Further Readings) that you would recommend to someone looking to read further on this topic.

Commented [A4R3]: This has already been done in the body of the chapter. So I haven't added this new section.

Discussion of the Literature (500-1,000 words) (**This does not need to be a separate section if this occurs throughout the manuscript)

Please discuss, briefly, the main threads in scholarship on your topic, including past approaches to the subject as well as research questions that remain or that are currently being pursued. While this cannot serve as an exhaustive historiography, it should discuss, in broad strokes, the scholarship on your topic.

Further Reading

Banaji, S., & Bhat, R. (2021). *Social media and hate* (1st ed.). Routledge. https://doi.org/10.4324/9781003083078

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.

Brock Jr., A. (2020). *Distributed Blackness: African American cybercultures*. NYU Press. https://www.degruyter.com/isbn/9781479811908

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15.

Daniels, J. (2008). *Race, civil rights, and hate speech in the digital era* (pp. 1–27). City University of New York. https://academicworks.cuny.edu/gc_pubs/193

Gandy, O. H. (2009). *Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage*. Ashgate.

Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, *60*(4), 3–26. https://doi.org/10.1177/0306396818823172

Monahan, T. (2008). Surveillance and inequality. *Surveillance & Society*, *5*(3), 11. https://doi.org/10.24908/ss.v5i3.3421

Nakamura, L. (1995). Race in/for cyberspace: Identity tourism and racial passing on the internet. *Works and Days*, *13*(1 & 2), 181–193.

Nelson A. Tu T. L. N. & Hines A. H. (2001). *Technicolor: race technology and everyday life*. New York University.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Steele, C. K. (2021). *Digital Black feminism*. NYU Press.

References

Alizadeh, T., Grubesic, T. H., & Helderop, E. (2017). Urban governance and big corporations in the digital economy: An investigation of socio-spatial implications of Google Fiber in Kansas City. *Telematics and Informatics*, *34*(7), 973–986. https://doi.org/10.1016/j.tele.2017.04.007

Amoore, L. (2008). *Risk and the war on terror*. Routledge.

Aouragh, M., & Chakravartty, P. (2016). Infrastructures of empire: Towards a critical geopolitics of media and information studies. *Media, Culture & Society*, *38*(4), 559–575. https://doi.org/10.1177/0163443716643007

Banaji, S. (2015). Behind the high-tech fetish: Children, work and media use across classes in India. *International Communication Gazette*, *77*(6), 519–532.

Banaji, S., & Bhat, R. (2021). *Social media and hate* (1st ed.). Routledge. https://doi.org/10.4324/9781003083078

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, *104*, 1–16.

Baynes, L. M. (2004). Deregulatory injustice and electronic redlining: The color of access to telecommunications. *Administrative Law Review*, *56*, 263.

Beller, J. (2021). *The world computer: Derivative conditions of racial capitalism*. Duke University Press.

Ben-David, A., & Fernández, A. M. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, *10*(0), Article 0.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity.

Berendt, B., & Preibusch, S. (2012). Exploring discrimination: A user-centric evaluation of discrimination-aware data mining. *2012 IEEE 12th International Conference on Data Mining Workshops*, 344–351. https://doi.org/10.1109/ICDMW.2012.109

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5454–5476. https://doi.org/10.18653/v1/2020.acl-main.485

boyd, danah, & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Brock Jr., A. (2020). *Distributed Blackness: African American cybercultures*. NYU Press; /z-wcorg/. https://www.degruyter.com/isbn/9781479811908

Buckingham, D. (2006). Defining digital literacy: What do young people need to know about digital media? *Nordic Journal of Digital Literacy*, *1*(4), 263–277. https://doi.org/10.18261/ISSN1891-943X-2006-04-03

Bullock, E. C. (2017). Only STEM can save us? Examining race, place, and STEM education as property. *Educational Studies*, *53*(6), 628–641. https://doi.org/10.1080/00131946.2017.1369082

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, *81*, 1–15.

Burnap, P., & Williams, M. L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy & Internet*, *7*(2), 223–242. https://doi.org/10.1002/poi3.85

Calders, T., & Verwer, S. (2010). Three naive Bayes approaches for discrimination-free classification. *Data Mining and Knowledge Discovery*, *21*(2), 277–292. https://doi.org/10.1007/s10618-010-0190-x

Cave, S., & Dihal, K. (2020). The whiteness of AI. *Philosophy & Technology*, *33*(4), 685–703. https://doi.org/10.1007/s13347-020-00415-6

Chakravartty, P. (2019). Media, race and the infrastructures of empire. In *Media and Society* (pp. 245–262). Bloomsbury Academic.

Chakravartty, P., & da Silva, D. F. (2012). Accumulation, dispossession, and debt: The racial logic of global capitalism—An introduction. *American Quarterly*, *64*(3), 361–385.

Chun, W. H. K. (2009). Race and/as technology. *Camera Obscura: Feminism, Culture, and Media Studies*, *24*(1 (70)), Number 1. /z-wcorg/.

Chun, W. H. K. (2019). Queering homophily. In W. H. K. Chun, F. Cramer, H. Steyerl, & C. Apprich (Eds.), *Pattern discrimination*. University of Minnesota Press.

Citron, D. K., & Norton, H. L. (2011). Intermediaries and hate speech: Fostering digital citizenship for our Information Age. *Boston University Law Review*, *91*, 1435.

Coleman, B. (2004). *The art of disappearance: Autobiography, race, and technology (Jean-Jacques Rousseau, Frederick Douglass, Norbert Wiener, Filippo Tommaso Marinetti, Italy)*. New York University.

Coleman, B. (2009). Race as technology. *Camera Obscura: Feminism, Culture, and Media Studies*, *24*(1 (70)), 177–207. https://doi.org/10.1215/02705346-2008-018

Condis, M. (2021). Playing at racism: White supremacist recruitment in online video game culture. In *White Supremacy and the American Media*. Routledge.

Corda, M. (2007). *Protezione della privacy e prevenzione della discriminazione nel data mining* [Università degli Studi di Pisa]. https://www.tesionline.it/tesi/scienze-matematiche%2C-fisiche-e-naturali/protezione-della-privacy-e-prevenzione-della-discriminazione-nel-data-mining/21139

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. MIT Press.

Couldry, N., & Mejias, U. (2019). *The costs of connection: How data are colonizing human life and appropriating It for capitalism*. Stanford University Press.

Crawford, K., & Schultz, J. (2014). Big data and due process: Toward a framework to redress predictive privacy harms. *Boston College Law Review*, *55*, 93.

Croll, A. (2012, October 17). New ethics for a new world—O'Reilly Radar. *O'Reilly Radar*. http://radar.oreilly.com/2012/10/new-ethics-for-a-new-world.html

Crooks, R. (2022). *Toward people's community control of technology: Race, access, and education* (pp. 1–37). Just Tech, Social Science Research Council. https://www.ssrc.org/publications/toward-peoples-community-control-of-technology-race-access-and-education/

Dailey, D., Bryne, A., Powell, A., Karaganis, J., & Chung, J. (2010). *Broadband adoption in low-income communities*. Social Science Research Council.

Daniels, J. (2008). *Race, civil rights, and hate speech in the digital era* (pp. 1–27). City University of New York. https://academicworks.cuny.edu/gc_pubs/193

Daniels, J. (2015). "My brain database doesn't see skin color": Color-blind racism in the technology industry and in theorizing the web. *American Behavioral Scientist*, *59*(11), 1377–1393. https://doi.org/10.1177/0002764215578728

Daniels, J., Lalone, N., Embrick, D. G., Wright, J. T., & Lukacs, A. (2012). Connecting extremist and mainstream expressions of white supremacy. In *Social exclusion, power, and video game play* (pp. 83–97). Lexington Press.

Dijck, J. van, Poell, T., & Waal, M. de. (2018). *The platform society: Public values in a connective world*. Oxford University Press.

Dwork, C. (2017). What's fair? *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1. https://doi.org/10.1145/3097983.3105807

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 214–226. https://doi.org/10.1145/2090236.2090255

Dworkin, R. M. (1977). *Taking rights seriously*. Harvard University Press.

Edelman, B. G., & Luca, M. (2014). *Digital discrimination: The case of Airbnb.com* (pp. 1–21). Harvard Business School. https://www.hbs.edu/ris/Publication%20Files/Airbnb_92dd6086-6e46-4eaf-9cea-60fe5ba3c596.pdf

Feldman, M., Friedler, S. A., Moeller, J., Scheidegger, C., & Venkatasubramanian, S. (2015). Certifying and removing disparate impact. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 259–268. https://doi.org/10.1145/2783258.2783311

Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys*, *51*(4), 85:1-85:30. https://doi.org/10.1145/3232676

Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2016). On the (im)possibility of fairness. *arXiv:1609.07236 [Cs, Stat]*. http://arxiv.org/abs/1609.07236

Friedline, T., & Chen, Z. (2021). Digital redlining and the fintech marketplace: Evidence from US zip codes. *Journal of Consumer Affairs*, *55*(2), 366–388. https://doi.org/10.1111/joca.12297

Friedline, T., Naraharisetti, S., & Weaver, A. (2020). Digital redlining: Poor rural communities' access to fintech and implications for financial inclusion. *Journal of Poverty*, *24*(5–6), 517–541. https://doi.org/10.1080/10875549.2019.1695162

Friedman, B. (1996). Value-sensitive design. *Interactions*, *3*(6), 16–23. https://doi.org/10.1145/242485.242493

Friedman, B. (1997). *Human values and the design of computer technology*. Cambridge University Press/CSLI.

Friedman, B., Brok, E., Roth, S. K., & Thomas, J. (1996). Minimizing bias in computer systems. *ACM SIGCHI Bulletin*, *28*(1), 48–51. https://doi.org/10.1145/249170.249184

Friedman, B., & Nissenbaum, H. (1993). Discerning bias in computer systems. *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems*, 141–142. https://doi.org/10.1145/259964.260152

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, *14*(3), 330–347. https://doi.org/10.1145/230538.230561

Galperin, H., Le, T. V., & Wyatt, K. (2021). Who gets access to fast broadband? Evidence from Los Angeles County. *Government Information Quarterly*, *38*(3), 101594. https://doi.org/10.1016/j.giq.2021.101594

Gandy, O. H. (1993). *The panoptic sort: A political economy of personal information*. Westview.

Gandy, O. H. (2009). *Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage*. Ashgate.

Gangadharan, S. P. (2012). Digital inclusion and data profiling. *First Monday*, *17*(5–7), n.p. https://doi.org/10.5210/fm.v17i5.3821

Gangadharan, S. P., & Niklas, J. (2019). Decentering technology in discourse on discrimination. *Information, Communication & Society*, *22*(7), 882–899. https://doi.org/10.1080/1369118X.2019.1593484

Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal*, *33*(4), 111–117.

Gelber, K. (2021). Differentiating hate speech: A systemic discrimination approach. *Critical Review of International Social and Political Philosophy*, *24*(4), 393–414. https://doi.org/10.1080/13698230.2019.1576006

Gelber, K., & McNamara, L. (2016). Evidencing the harms of hate speech. *Social Identities*, *22*(3), 324–341. https://doi.org/10.1080/13504630.2015.1128810

Gillespie, T. (2010). The politics of "platforms." *New Media & Society*, *12*(3), 347–364. https://doi.org/10.1177/1461444809342738

Gilliard, C. (2017, July 3). Pedagogy and the logic of platforms. *Educause*. https://er.educause.edu/articles/2017/7/pedagogy-and-the-logic-of-platforms

Gilliom, J. (2001). *Overseers of the poor: Surveillance, resistance, and the limits of privacy*. University of Chicago Press.

Gill-Peterson, J. (2014). The Technical Capacities of the Body: Assembling Race, Technology, and Transgender. *TSQ: Transgender Studies Quarterly*, *1*(3), 402–418. https://doi.org/10.1215/23289252-2685660

Gonzalez, H. B., & Kuenzi, J. J. (2012, August 1). *Science, technology, engineering, and mathematics (STEM) education: A Primer* (United States) [Report]. UNT Digital Library; Library of Congress. Congressional Research Service. https://digital.library.unt.edu/ark:/67531/metadc122233/

Gonzalez, J. (2006). Morphologies: Race as a visual technology. In M. Flanagan & A. Booth (Eds.), *Re: Skin* (pp. 339–353). MIT Press.

Gray, K. L. (2012). Deviant bodies, stigmatized identities, and racist acts: Examining the experiences of African-American gamers in Xbox Live. *New Review of Hypermedia and Multimedia*, *18*(4), 261–276. https://doi.org/10.1080/13614568.2012.746740

Hanna, A., Denton, E., Smart, A., & Smith-Loud, J. (2020). Towards a critical race methodology in algorithmic fairness. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 501–512. https://doi.org/10.1145/3351095.3372826

Harcourt, B. E. (2007). *Against prediction: Profiling, policing, and punishing in an actuarial age*. University of Chicago Press.

Hargittai, E. (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4). Scopus. https://doi.org/10.5210/fm.v7i4.942

Helberger, N., Pierson, J., & Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, *34*(1), 1–14. https://doi.org/10.1080/01972243.2017.1391913

Hervik, P. (2019). Racialization in the Nordic countries: An introduction. In P. Hervik (Ed.), *Racialization, racism, and anti-racism in the Nordic countries* (pp. 3–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-74630-2_1

Hobbs, R. (2020). Propaganda in an age of algorithmic personalization: Expanding literacy research and practice. *Reading Research Quarterly*, *55*(3), 521–533. https://doi.org/10.1002/rrq.301

Hochman, B. (2014). *Savage preservation: The ethnographic origins of modern media technology*. University of Minnesota Press.

Hoffmann, A. L. (2019). Where fairness fails: Data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, *22*(7), 900–915. https://doi.org/10.1080/1369118X.2019.1573912

Howard, B. (1995, July 1). Will there be redlining in cyberspace? Lack of access may trap African American in information ghettos. *Black Enterprise*.

Hutchinson, B., & Mitchell, M. (2019). 50 years of test (un)fairness: Lessons for machine learning. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 49–58. https://doi.org/10.1145/3287560.3287600

Iginio, G., Danit, G., Thiago, A., & Gabriela, M. (2015). *Countering online hate speech*. UNESCO Publishing.

Johnson, P. (2016). Racial technologies in the time of black cyborgnetic consciousness. In *The Routledge Companion to Biology in Art and Architecture*. Routledge.

Kahl, C. M. (1997). Electronic redlining: Racism on the information superhighway? *The Katharine Sharp Review*, *4*, 1–8.

Kamiran, F., & Calders, T. (2009). Classifying without discriminating. *2009 2nd International Conference on Computer, Control and Communication*, 1–6. https://doi.org/10.1109/IC4.2009.4909197

Katz, Y. (2020). *Artificial Whiteness: Politics and ideology in artificial intelligence*. Columbia University Press.

Keeling, K. (2014). Queer OS. *Cinema Journal*, *53*(2), 152–157. https://doi.org/10.1353/cj.2014.0004

Keipi, T., Näsi, M., Oksanen, A., & Räsänen, P. (2017). *Online hate and harmful content: Cross-national perspectives*. Routledge. https://library.oapen.org/viewer/web/viewer.html?file=/bitstream/handle/20.500.12657/22350/9781138645066_text.pdf?sequence=1&isAllowed=y

Kenyon, A. T., & Scott, A. (Eds.). (2020). *Positive free speech*. Hart Publishing.

Klawitter, E., & Hargittai, Eszter. (2018). "It's like learning a whole other language:" The role of algorithmic skills in the curation of creative goods. *International Journal of Communication*, *12*, 3490–3510. Scopus.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J., Robinson, D., & Yu, H. (2015). Accountable algorithms. *University of Pennsylvania Law Review*, *165*, 633–705.

Kwet, M. (2019). Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, *60*(4), 3–26. https://doi.org/10.1177/0306396818823172

Lambright, K. (2019). Digital redlining: The Nextdoor App and the neighborhood of make-believe. *Cultural Critique*, *103*, 84–90. https://doi.org/10.5749/culturalcritique.103.2019.0084

Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review*, *7*, 3–14.

Livingstone, S., & Helsper, E. (2007). Gradations in digital inclusion: Children, young people and the digital divide. *New Media and Society*, *9*(4), 671–696. Scopus. https://doi.org/10.1177/1461444807080335

Lyon, D. (1994). *The electronic eye: The rise of surveillance society*. Polity.

Marjanovic, O., Cecez-Kecmanovic, D., & Vidgen, R. (2021). Theorising algorithmic justice. *European Journal of Information Systems*, 1–19. https://doi.org/10.1080/0960085X.2021.1934130

Matamoros-Fernández, A. (2017). Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society*, *20*(6), 930–946. https://doi.org/10.1080/1369118X.2017.1293130

Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*. https://doi.org/10.1177/1527476420982230

McMillan Cottom, T. (2020). Where platform capitalism and racial capitalism meet: The sociology of race and racism in the digital society. *Sociology of Race and Ethnicity*, *6*(4), 441–449. https://doi.org/10.1177/2332649220949473

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, *33*(4), 659–684. https://doi.org/10.1007/s13347-020-00405-8

Monahan, T. (2008). Surveillance and inequality. *Surveillance & Society*, *5*(3), 11. https://doi.org/10.24908/ss.v5i3.3421

Morales-Doyle, D., & Gutstein, E. “Rico.” (2019). Racial capitalism and STEM education in Chicago Public Schools. *Race Ethnicity and Education*, *22*(4), 525–544. https://doi.org/10.1080/13613324.2019.1592840

Morozov, E. (2011). *The Net delusion: The dark side of Internet freedom*. PublicAffairs. http://www.publicaffairsbooks.com/morozovch1.pdf

Mulligan, D. K., & Dwork, C. (2013). It’s not privacy, and it’s not fair. *Stanford Law Review*, *66*(35), 35–40. March 3, 2017.

Nadaf, A. H. (2021). “Lockdown within a lockdown": The “digital redlining” and paralyzed online teaching during COVID-19 in Kashmir, a conflict territory. *Communication, Culture and Critique*, *14*(2), 343–346. https://doi.org/10.1093/ccc/tcab019

Nakamura, L. (1995). Race in/for cyberspace: Identity tourism and racial passing on the internet. *Works and Days*, *13*(1 & 2), 181–193.

Nakamura, L. (2007). *Digitizing race: Visual cultures of the internet*. University of Minnesota Press.

Nakamura, L. (2017). Racism, sexism, and gaming’s cruel optimism. In J. Malkowski & T. M. Russworm (Eds.), *Gaming Representation: Race, Gender, and Sexuality in Video Games* (pp. 245–250). Indiana University Press.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press; /z-wcorg/.

Noble, S. U., & Roberts, S. T. (2019). Technological elites, the meritocracy, and postracial myths in Silicon Valley. In R. Mukherjee, S. Banet-Weiser, & H. Gray (Eds.), *Racism postrace* (p. 0). Duke University Press. https://doi.org/10.1215/9781478003250-008

Nopper, T. K. (2019). Digital character in “The scored society:” FICO, social networks, and competing measurements of creditworthiness. In R. Benjamin (Ed.), *Captivating*

technology: Race, carceral technoscience, and liberatory imagination in everyday life (pp. 170–187). Duke University Press. https://doi.org/10.1515/9781478004493-009

Pasquale, F. (2015). The black box society: The secret algorithms that control money and information. In *The Black Box Society*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Pedreschi, D., Ruggieri, S., & Turini, F. (2009). Integrating induction and deduction for finding evidence of discrimination. *Proceedings of the 12th International Conference on Artificial Intelligence and Law*, 157–166. https://doi.org/10.1145/1568234.1568252

Pedreschi, D., Ruggieri, S., & Turini, F. (2008). Discrimination-aware data mining. *ACM Proceedings*. KDD '08, Las Vegas, Nevada. https://doi.org/10.1145/1401890.1401959

Plasencia, M. M. (1999). The Politics of Race on the Electronic Highway: An Analysis of the Video Dialtone Redlining Cases, and the Nynex Consent Decree in Roxbury. *Touro Law Review*, *15*(2), 17.

Pohjonen, M., & Udupa, S. (2017). Extreme speech online: An anthropological critique of hate speech debates. *International Journal of Communication*, *11*(0), Article 0.

Poster, W. R. (2019). Racialized surveillance in the digital service economy: Race, carceral technoscience, and liberatory imagination in everyday life. In R. Benjamin (Ed.), *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life* (pp. 133–169). Duke University Press. https://doi.org/10.1515/9781478004493-009

Powell, C. (2018). Race and Rights in the Digital Age. *AJIL Unbound*, *112*, 339–343. Cambridge Core. https://doi.org/10.1017/aju.2018.89

Prieger, J. E., & Hu, W.-M. (2008). The broadband digital divide and the nexus of race, competition, and quality. *Information Economics and Policy*, *20*(2), 150–167. https://doi.org/10.1016/j.infoecopol.2008.01.001

Raengo, A. (2013). *On the sleeve of the visual: Race as face value*. Dartmouth College Press.

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). Saving face: Investigating the ethical concerns of facial recognition auditing. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 145–151. https://doi.org/10.1145/3375627.3375820

Randolph, L. B. (1995, January 1). Blacks in the fast lane of the information superhighway. *Ebony*.

Regan, P. M. (1995). *Legislating privacy: Technology, social values, and public policy*. University of North Carolina Press; /z-wcorg/.

Roberts, D. E. (2009). Race, gender, and genetic Technologies: A new reproductive dystopia? *Signs: Journal of Women in Culture and Society*, *34*(4), 783–804. https://doi.org/10.1086/597132

Sambasivan, N., Arnesen, E., Hutchinson, B., Doshi, T., & Prabhakaran, V. (2021). Re-imagining Algorithmic Fairness in India and Beyond. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 315–328. https://doi.org/10.1145/3442188.3445896

Sánchez-Monedero, J., Dencik, L., & Edwards, L. (2020). What does it mean to "solve" the problem of discrimination in hiring? Social, technical and legal perspectives from the UK on automated hiring systems. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 458–468. https://doi.org/10.1145/3351095.3372849

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014). *Auditing algorithms: Research methods for detecting discrimination on internet platforms*. New America.

Scannell, R. J. (2019). This is not minority report: Predictive policing and population racism. In R. Benjamin (Ed.), *Captivating technology: Race, carceral technoscience, and liberatory imagination in everyday life* (pp. 107–130). Duke University Press. https://doi.org/10.1515/9781478004493-009

Schauer, F. F. (2003). *Profiles, probabilities, and stereotypes*. Belknap Press of Harvard University Press.

Schmidt, A., & Wiegand, M. (2019). *A Survey on Hate Speech Detection using Natural Language Processing*. 1–10. https://ids-pub.bsz-bw.de/frontdoor/index/index/docId/8426

Selbst, A. D., boyd, danah, Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. https://doi.org/10.1145/3287560.3287598

Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, *6*, 341–362.

Shin, D., Rasul, A., & Fotiadis, A. (2021). Why am I seeing this? Deconstructing algorithm literacy through the lens of users. *Internet Research*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/INTR-02-2021-0087

Sinclair, B. (Ed.). (2004). *Technology and the African-American experience: Needs and opportunities for study*. MIT Press.

Steele, C. K. (2021). *Digital Black feminism*. NYU Press.

Stevens, N., & Keyes, O. (2021). Seeing infrastructure: Race, facial recognition, and the politics of data. *Cultural Studies*, *35*(4–5), 833–853. https://doi.org/10.1080/09502386.2021.1895252

Sweeney, L. (2013). Discrimination in online ad delivery. *arXiv:1301.6822 [Cs]*. http://arxiv.org/abs/1301.6822

Taborn, T. D. (2008). Separating race from technology: Finding tomorrow's IT progress in the past. In A. Everett (Ed.), *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning* (pp. 39–59). MIT Press.

Taylor, A., & Sadowski, J. (2015, May 27). Digital redlining. *The Nation*, *300*(24), 24–27.

Taylor, L. (2017). What is data justice: The case for connecting digital rights and freedoms globally. *Big Data & Society*, *4*(2), 1–14.

Thatcher, J., O'Sullivan, D., & Mahmoudi, D. (2016). Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space*, *34*(6), 990–1006. https://doi.org/10.1177/0263775816633195

Udupa, S., Gagliardone, I., & Hervik, P. (2021). *Digital hate: The global conjuncture of extreme speech*. Indiana University Press. https://publish.iupress.indiana.edu/projects/digital-hate

U.S. Executive Office of the President. (2014). *Big data: Seizing opportunities, preserving values* (pp. 1–85). U.S. Executive Office of the President. https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale (ISS). *Information, Communication & Society*, *19*(6), 804–823. https://doi.org/10.1080/1369118X.2015.1078834

Van Dijk, J., & Hacker, K. (2003). The digital divide as a complex and dynamic phenomenon. *Information Society*, *19*(4), 315–326. Scopus. https://doi.org/10.1080/01972240309487

Venkatasubramanian, S. (2019). Algorithmic fairness: Measures, methods, and representations. *Proceedings of the 38th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, 481. https://doi.org/10.1145/3294052.3322192

Warschauer, M. (2002). Reconceptualizing the digital divide. *First Monday*, *7*. https://doi.org/10.5210/fm.v7i7.967

Winner, L. (1986). *The whale and the reactor*. Chicago University Press.

Zarouali, B., Helberger, N., & Vreese, C. H. de. (2021). Investigating algorithmic misconceptions in a media context: Source of a new digital divide? *Media and Communication*, *9*(4), Article 4.