Received: 19 June 2021 | Revised: 19 January 2022 | Accepted: 19 January 2022

DOI: 10.1111/soc4.12962

ARTICLE

WILEY

Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates

Mike Zajko

Department of History and Sociology, The University of British Columbia Okanagan, Kelowna, British Columbia, Canada

Correspondence
Mike Zajko, Department of History and Sociology, The University of British Columbia Okanagan, Kelowna, British Columbia, V1V 1V7, Canada.
Email: mike.zajko@ubc.ca

Funding information
University of British Columbia; Social Sciences and Humanities Research Council of Canada, Grant/Award Number: 430-2021-00810

Abstract

Artificial intelligence (AI) and algorithmic systems have been criticized for perpetuating bias, unjust discrimination, and contributing to inequality. Artificial intelligence researchers have remained largely oblivious to existing scholarship on social inequality, but a growing number of sociologists are now addressing the social transformations brought about by AI. Where bias is typically presented as an undesirable characteristic that can be removed from AI systems, engaging with social inequality scholarship leads us to consider how these technologies reproduce existing hierarchies and the positive visions we can work towards. I argue that sociologists can help assert agency over new technologies through three kinds of actions: (1) critique and the politics of refusal; (2) fighting inequality through technology; and (3) governance of algorithms. As we become increasingly dependent on AI and automated systems, the dangers of further entrenching or amplifying social inequalities have been well documented, particularly with the growing adoption of these systems by government agencies. However, public policy also presents some opportunities to restructure social dynamics in a positive direction, as long as we can articulate what we are trying to achieve, and are aware of the risks and limitations of utilizing these new technologies to address social problems.

KEYWORDS
algorithms, artificial intelligence, bias, inequality, public policy, technology

1 | INTRODUCTION

Over the past several years, the idea that society is being radically transformed through artificial intelligence (AI), machine learning, and automation has been widely discussed in popular venues. This transformation has been addressed through a growing body of sociological scholarship (see Joyce et al., 2021; also; Airoldi, 2022; Burrell & Fourcade, 2021; Davis et al., 2021; Elliott, 2022; Issar & Aneesh, 2022; Jaton, 2021), although some of the most prominent sociological analyses of these technologies have come from academics who are not (formally) sociologists (Crawford, 2021; Eubanks, 2017; Zuboff, 2018). Societal questions are now recognized as crucial ones for the developers of new algorithmic technologies, but 'AI scientists continue to demonstrate a limited understanding of the social' (Joyce et al., 2021, p. 5). Much of the discourse addressing algorithmic technologies remains characterized by a sense of inevitability and technological determinism (Vicsek, 2020), and one of the biggest concerns around the 'rise of AI' has been a corresponding loss of human agency (Anderson & Rainie, 2018). This article will sketch out some of the ways that sociologists can actively contribute to a better world, or oppose the dangerous visions that are so common in this space.

I am particularly interested in how sociologists can approach a set of problems in the field of AI, machine learning, and automated decision-making (ADM) that are predominantly understood as issues of 'bias' by technologists, but which have roots in pre-existing social inequalities. Given the ongoing proliferation of algorithmic systems, sociologists can study these technologies and their unequal consequences in familiar domains such as work, education, and law. However, to go beyond an empirical analysis of current developments, there are three kinds of contributions sociologists can make through praxis, policy influence, or interdisciplinary collaboration: (1) critique and the politics of refusal; (2) fighting inequality through technology; and (3) the governance of algorithmic governance. All three should be considered important modes of positive action at a time when societies are being reconfigured through algorithmic governance, and when the algorithmic reproduction of inequality is increasingly recognized as a major issue. Critique and the politics of refusal represent a well-worn path for sociologists, through which we can also articulate a positive vision for the world, even if our agency to pursue it remains limited. For more direct influence, there are opportunities for sociologists to participate in the design and implementation of algorithmic systems to improve social outcomes, but our participation can carry the risk of strengthening already-dominant interests. One of the most important interdisciplinary contributions from sociology may be to unsettle and reconstitute established problems in new ways.

2 | BACKGROUND

In much of the world, automated and algorithmic decision-making systems now govern various aspects of people's lives, including those most directly implicated in social inequality, such as employment, government benefits, and criminalization (Benjamin, 2019; Eubanks, 2017; O'Neil, 2016; Schuilenburg & Peeters, 2021). It is now increasingly recognized that machine systems built to be objective and unbiased do indeed discriminate along familiar human lines, reproducing or amplifying social differences and inequalities (Noble, 2018). The disposition of automated systems to uphold the existing social order has been described as a 'conservative' tendency (Birhane, 2020; Zajko, 2021), but among AI researchers and developers, the general understanding is that people (or society) exhibit various 'biases' which are then reproduced in automated systems. This is particularly the case for today's dominant forms of AI or 'machine learning' algorithms, which must be 'trained' using datasets that reflect human judgments, priorities, and conceptual categories (Crawford, 2021). When this data is biased, or when the underlying 'ground truth' of society in unequal, this inequality can be encoded and reproduced in algorithms based on machine learning (Hacker, 2018). Data about patterns in society serves as the input on which these systems are trained, with the resulting automated decisions (the output) reflecting and perpetuating social inequalities. Machine learning and algorithmic governance therefore frequently create 'feedback loops' that replicate and amplify existing patterns in society (Brayne, 2020; Brayne & Christin, 2021; Mehrabi et al., 2019).

The solution to this problem among AI developers has been to find various ways to remove, reduce or 'minimize' bias in datasets and algorithmic decisions (Silberg & Manyika, 2019). This is now increasingly recognized to be a very complex and multi-dimensional challenge that cannot be achieved purely through technological solutions (Kind, 2020). None of this has been surprising to scholars in fields such as science and technology studies (STS), critical data studies, or critical algorithm studies, who have long examined the relationships between technology and social structures (Felt et al., 2017; Gillespie & Seaver, 2016; Iliadis & Russo, 2016). Sociologists in general have been less engaged in these issues thus far, which is unfortunate given that social theory can fill the conceptual chasm currently identified as 'bias' in data science (Joyce et al., 2021; Zajko, 2021).

While some forms of bias in AI are roughly equivalent to how this term is understood in social research methods and sampling (particularly given the statistical foundations for today's AI technologies), other biases in AI reflect what would be understood by a sociologist as various kinds of inequality. These inequalities are sometimes categorized as 'societal bias' (Friedman & Nissenbaum, 1996) or 'historical bias' (Suresh & Guttag, 2020) by AI practitioners, which has the effect of black-boxing any specifics about where the bias comes from or is structured. But sociologists can name the intersecting structures (Davis et al., 2021; Hoffmann, 2019) of inequality, oppression, or exploitation that produced the biases AI and data science practitioners are concerned about: colonialism, heteropatriarchy, ableism, white supremacy, and various articulations of capitalism or political economy. While AI developers have been repeatedly criticized for being unwilling to look beyond their own field to 'see the social' (Irani & Chowdhury, 2019), this has slowly been changing, and social scientists have been identified as a source of relevant expertise by AI researchers (CIFAR, 2020; Kusner & Loftus, 2020), tech companies (Maris, 2022), academic bodies and governments (G7 Science Academies, 2019). In short, there is an opportunity for sociologists to address pressing issues concerning the automation of inequality, by building on the field's long history of studying social stratification, power, oppression, and intersectionality (Joyce et al., 2021).

3 | BIAS AS SOCIETY'S "UNEQUAL GROUND TRUTH"

In sociology, teaching students that society is 'not fair' is a regular part of the undergraduate curriculum. Media and cultural studies have long documented inequalities in how different groups of people are represented, the exclusion of groups, and how media privilege dominant categories (e.g., Erigha, 2015; Pascale, 2013). For developers of AI systems however, the evident unfairness of algorithmic technologies (including blatantly racist or sexist algorithmic outputs) has in recent years created major conceptual difficulties for which they were untrained and unprepared. Solutions have been developed to minimize bias or increase fairness, but there is no consensus on what a fair or unbiased algorithm might look like, or if this is even achievable (Green & Hu, 2018; Silberg & Manyika, 2019). Artificial intelligence developers refer to the reality that exists outside of their models as the 'ground truth', and bias is often defined as deviations from this truth, or inaccurate representations and predictions. But when the truth is that society is deeply, structurally unjust and unequal, and that technologies are part of these structures, the question is whether our algorithms should accurately reproduce inequality or work to change it.

While the conventional approach in AI and data science treats bias as 'error' – inaccurate modeling or prediction – a broader definition of bias that is also used (but less often articulated) equates it with whatever might be undesirable in society (Mitchell et al., 2020; Zajko, 2021). In short, while an algorithmic model that accurately reproduces society's 'unequal ground truth' might be considered unbiased in the first narrow sense of bias as accuracy, it may be biased according to this second view of bias as an undesirable tendency. The problem for data science is that while accuracy presents a clear goal to work towards through the elimination or reduction in bias, the second definition of bias as an undesirable tendency does not specify what is desirable, beyond the contested notions of equality and fairness (both of which have a number of competing and irreconcilable definitions, see Friedler et al., 2016; Silberg & Manyika, 2019).

Essentially, what has happened in recent years is that AI scholarship has 'discovered' that social inequality is real. This inequality consistently shows up in the data used to train algorithms through machine learning, as well as the outputs of these systems. Practitioners have approached these problems through the language of bias, but bias is not helpful in understanding how to deal with pre-existing inequalities that do not result from flaws in data or methods. While there are many inequalities that are deemed to be generally problematic or undesirable, there is far less agreement on what a desirable world looks like, and the aversion that many technologists feel towards taking a political stance has also limited this discussion. Google, for instance, has explained offensive search results as accurate representations of a ground truth (either the search behaviour of other users, or of online content – see Gibbs, 2016). In situations where Google has modified search algorithms producing racist and sexist outputs, it has often done so quietly, in response to media and public attention over specific cases (Noble, 2018), pressure from external interests, or for internal reasons (West et al., 2019).

Whether we are looking at algorithmic search results, hiring decisions, or recidivism prediction, we will continue to confront examples of harmful discrimination that cannot be reduced to issues of accuracy; any attempt at a solution will be inherently political, and avoiding the issue or treating it as being 'out of scope' for data science amounts to a conservative orientation (Green, 2018; Zajko, 2021). As a consequence, a growing movement among AI researchers has been taking a more radical path to altering relations of power (Kalluri, 2020), but interdisciplinary contributions from scholars of power and inequality remain badly needed.

4 | SOCIAL INEQUALITY AND TECHNOLOGY: THE VIEW FROM SOCIOLOGY

Sociologists presume the existence of inequalities in any domain of society, and recognize that the benefits and harms of technology are not evenly distributed. Scholarship on the 'digital divide' has examined unequal access and use of technologies since the 1990s (van Dijk, 2006; Wellman & Haythornthwaite, 2002); when the Internet was presented as a broadly empowering force for social good, sociologists warned that these technologies would also reinforce numerous existing inequalities (DiMaggio et al., 2004). Selwyn (2004) and Helsper (2008, 2012) drew on Bourdieusian theory and the concept of 'social exclusion' (see also van Dijk, 2005), to redefine the problem of the digital divide as one of 'digital exclusion', with socio-technical inequalities 'rooted in broader social categories linked to other types of disadvantage and discrimination' (Helsper, 2012, p. 406). Park and Humphry (2019) extended this focus on exclusion to AI and ADM, analyzing how automated systems discriminate, deny, and punish welfare recipients in Australia (see also Eubanks, 2017; James & Whelan, 2021; Schou & Pors, 2019). Relatedly, sociologists of education (Davies et al., 2021; Williamson & Eynon, 2020) and medical technologies (Roberts & Rollins, 2020; Singh & Steeves, 2020) have studied how algorithms exclude and discriminate on the basis of pre-existing social inequalities.

Sociologists of work, occupations, and organizations have examined how new algorithmic systems reconfigure human labor, in ways that are often detrimental to workers (Bailey et al., 2020; Griesbach et al., 2019; Kellogg et al., 2020; Newlands, 2021; Shestakofsky, 2017). Utopian and dystopian predictions of robots taking and transforming human jobs have been the subject of discourse analysis (see James & Whelan, 2021; Ossewaarde & Gulenc, 2020; Vicsek, 2020), but sociological scholarship has been skeptical or critical of these claims, and more attentive to questions of power relations (Boyd & Holton, 2018). Judy Wajcman contextualizes the latest wave of 'hyperbole about AI' (2017, p. 121) within 'the perennial anxiety about automation' (p. 125), reminding us to remain focused on the concentration of power with a few large corporations and the biases of Silicon Valley's (typically male and white) engineers. With the benefit of some historical perspective, our worries about the looming automation of human labor is misplaced, when 'the real issue is the unequal distribution of work, time and money that exist already' and how new technologies are creating 'not less work but more worse jobs' (Wajcman, 2017, p. 124).

The inequalities that are reproduced and reshaped through algorithmic technologies can be studied within organizations, but inequalities also play out on a global scale (Sampath, 2021), including international labor (Aneesh, 2009), and the flow of capital through colonial and extractive processes (Couldry & Mejias, 2019). Sociology's own coloniality (Bhambra, 2014) has often excluded societies of the 'Global South' from analysis, or treated societies of the North as universal (Go, 2020; Milan & Treré, 2019). While the most technologically-developed industrial nations in North America, Europe, and East Asia are presented as the key competitors in the 'race for AI' (Walch, 2020), algorithmic systems reach around the world and depend on global resources (Crawford, 2021). In the Global South, AI systems have been promoted for international development (Sinanan & McNamara, 2021), but are often complicit in processes of colonization and extraction (Couldry & Mejias, 2019; Kwet, 2019; Mohamed et al., 2020). Populations in the Global South have a different relationship with major AI platforms than those who live and work where these companies are headquartered (Aneesh, 2009; Gray & Suri, 2019; Grohmann & Araújo, 2021). Ideas from the Global South (and Indigenous epistemologies in Northern societies) can also be the source of conceptual alternatives to dominant theories and ethical rationalities of AI (Milan & Treré, 2019), including new ways to theorize, analyze, and critique these socio-technical systems (Birhane, 2021; Lewis et al., 2020). Sociologists can play a valuable role in bringing forward perspectives from social positions that have historically been marginalized in the development of AI.

5 | THE FUTURE OF INEQUALITY AND SOCIOLOGY'S RESPONSE

As the previous section outlined, numerous sociologists are working in domains where algorithmic and AI systems are impossible to ignore, and carry significant consequences. Some sociologists have addressed AI in a general sense, either to argue that it is indeed transforming our everyday lives (Elliott, 2019), or to push back against grandiose claims of AI's capabilities (Collins, 2018). Burrell and Fourcade (2021) remind us that 'we can both reject magical thinking about machine intelligence and acknowledge the enormous economic, political, and cultural power of the tech industry to transform the world we live in' (p. 231). Questions of power (see Kalluri, 2020) should be a primary concern for us as sociologists, given what is at stake in this ongoing transformation and the inequalities that algorithmic systems can either shift or lock into place.

Projecting from current trends, it is easy to imagine a future world where wealth and power are even more concentrated in the hands of the 'coding elite' (Burrell & Fourcade, 2021) than today – a scenario that one RAND-affiliated author has called 'Bezos World' (Lempert, 2019). This sort of speculative dystopia has already generated social engineering proposals to redistribute wealth and inequality along both radical and conservative lines – from 'luxury communism' (Bastani, 2019), to new capital taxes that fund basic income for citizens (Clifford, 2021), or a 'cyber republic' in which new kinds of ownership and 'data property rights' are maintained through tokens and distributed ledgers (Zarkadakis, 2020). These scenarios are typically presented as developing from an economic or political crisis emerging as a consequence of widespread automation, whether this is imagined in the near future or as currently underway. However, the social (re)engineering of inequality need not be contingent on some 'new Industrial Revolution' where robots replace human labor. Automated systems already distribute various goods and effects unequally, or are complicit in the reproduction of long-established inequalities. The future, in this sense, promises more of the same; control over sociotechnical systems will be used to further concentrate power along already-dominant lines.

Sociologists can theorize these developments by examining how social inequalities are structured, highlighting political economy, capitalism, and colonial relations (Couldry & Mejias, 2019; Dyer-Witheford et al., 2019;

Shestakofsky, 2020). While macro-level social theories provide analytic tools for global transformations, sociologists can attend to the production of power and knowledge through genealogies (Denton et al., 2021) and ethnographies of AI research (Hoffman, 2021; Jaton, 2021). There will also be continuing value in producing ethnographies (and institutional ethnographies, James & Whelan, 2021) of organizations implementing algorithmic systems (Bailey et al., 2020; Brayne & Christin, 2021; Cruz, 2020; Shestakofsky & Kelkar, 2020), as well as studies into the experiences of people who are further 'downstream', interacting with algorithmic systems (Christin, 2020; Noble, 2018). Roberge and Castelle (2021) argue that we need an 'end-to-end sociology' of AI, investigating how these 'upstream' and 'downstream' processes are entangled, and tracing these sociotechnical systems 'from genesis to impact and back again' (Roberge & Castelle, 2021, p. 3). Any analysis that takes a wider approach to the sociotechnical systems and their relations also needs to consider the discourses and imaginary futures that are motivating these developments, including the promotion of ideas about what kinds of futures are desirable or inevitable (James & Whelan, 2021; Schuilenburg & Peeters, 2021; Vicsek, 2020).

Given that AI and algorithmic systems can be studied sociologically, how can sociology contribute to current debates about these technologies? If the current trajectory leads to greater concentrations of wealth and power and the reproduction of existing inequalities, how might we interrupt these developments? The loss of human agency is a recurring worry in discussions of automation and AI (Anderson & Rainie, 2018), but we should recognize that structured inequality has long deprived marginalized groups of certain forms of agency (Shah, 2018). The following sections explain the different roles that sociologists can play in exercising our agency to change social structures and shape a better world, by addressing algorithmically-mediated inequalities.

5.1 | Critique

Critique is a valuable mode of sociological argument that can analyze 'how algorithms reproduce and reinforce existing structures', situating new technologies within 'broader political, racial, cultural, and economic formations' (Christin, 2020, p. 900). In this regard, Ruha Benjamin's (2019) work has been exemplary (preceded notably by Virginia Eubanks [2017] and Noble [2018]), as is that of many more interdisciplinary scholars in STS and related fields (Gillespie & Seaver, 2016; Hoffmann, 2019). Critique helps unpack the politics of algorithmic technology, perhaps as inspired by Winner (1980), or more contemporary examples from non-sociologists who engage with social theory (Crawford, 2021; Zuboff, 2018). It is work that remains badly needed in this space, particularly in a form that is understandable by technologists for whom the language of power, politics, and 'social impacts' remains unfamiliar terrain (CIFAR, 2020). Although writing across disciplines presents its own challenges within the field of sociology, critique is an area of strength for us and can draw on many existing theories, skills, and methods.

While understanding the technical sophistication of AI can seem daunting if this is considered a prerequisite for scholarly engagement, a critique of these systems can often begin by stating the sociologically obvious. Artificial intelligence practitioners are continuously applying statistical methods to categorizing human populations, with very little understanding of the social categories being operationalized. One recent study (since retracted due to lack of ethics approval) cited 19th century criminologist Cesare Lombroso as a justification for identifying criminality through facial features (Hashemi & Hall, 2020). Artificial intelligence practitioners continue to regularly produce work that can be characterized as physiognomy or phrenology (Stark & Hutson, forthcoming; Stinson, 2021), classifying individuals on the basis of discrete races, genders, and emotional states, through only the shallowest ontological engagement with these phenomena (Barrett et al., 2019; Hanna et al., 2019; Scheuerman et al., 2021). This work typically demonstrates little-to-no awareness of the pre-AI history of 'social sorting' (Lyon, 2003), and an inability to situate these systems in contemporary relations of governance or capital. Even though Silicon Valley-firms have employed significant numbers of social scientists, humanities scholars, and even voices critical of the industry, the possibilities for critique within these organizations are tenuous or circumscribed – as evidenced by Timnit Gebru's ouster from Google's AI

Ethics team and the events that have followed (Grant et al., 2021; Simonite, 2021). Independent academic critique remains badly needed from a variety of perspectives.

Critiques have sometimes been dismissed (particularly by practitioners) as having nothing to offer in terms of recommendations or preferred policies. Many sociological critiques have indeed limited themselves to 'unmasking hidden structures of inequality' (Cancian, 1995, p. 347), and stopped short of promoting social change or addressing themselves to policymakers. But critique, rejection, and opposition can be the basis for various kinds of normative arguments and practical ways forward. The 'politics of refusal' (Simpson, 2014) is at its core, a call to both oppose what is wrong and to do things differently (see Cifor et al., 2019). As applied to algorithmic systems and AI (Crawford, 2021; Eubanks, 2017), the politics of refusal is a more radical position than attempts to steer technologies in a more desirable direction, and it entails more than just refusing to contribute to the development of these systems. Rather, the politics of refusal takes a normative stand against some aspects of algorithmic control, automation, data extraction, and AI development. This opposition may be directed against specific applications of AI, such as facial recognition and physiognomy (Stark & Hutson, forthcoming) or predictive policing (Roberts, 2019). Refusal of algorithmic technologies may also be one part of the larger project of dismantling unjust and violent institutions, such as the 'carceral state' (Roberts, 2019), 'digital poorhouse' (Eubanks, 2017), and colonial structures (Walter et al., 2020). Arguments for refusal are sometimes part of a more general critique of technological determinism and inevitability (Benjamin, 2019; Crawford, 2021). Technologists and a broader public need to be regularly reminded that just because we *can* build or automate something, does not mean that we *should*, and that rejection of the latest 'innovation' in social control is a perfectly reasonable option.

Like the alternative approaches outlined below, the politics of refusal can be considered an exercise of human agency over sociotechnical systems, and like these alternative approaches it should also be understood in positive terms. As theorized by Audra Simpson (2014) and those who have built on her work, refusal is 'generative' (see McGranahan, 2016) and a means to affirm who we are, our relations, and our values. Arguing that there are certain decisions which must not be automated, contains within it a positive argument for what certain decisions should entail, in terms of human involvement. Because of this, arguing that we should *not* build or implement a system is valuable and often appropriate as a way to achieve desirable futures.

5.2 | Fighting inequality through technology

Structured inequality is reproduced on an ongoing basis and resistant to change, but the behaviour of digital systems is written in code, and code can be re-written. In other words, when inequalities are encoded in algorithms, software, or spreadsheets, changing the code will redistribute the inequalities. Research that addresses fairness and bias in data science is already trying to change the world by reducing bias, and algorithmic governance is delivering new ways to configure, engineer, and structure society. It is necessary to reject or resist some of these developments, but we can also try to steer them away from their darkest possibilities, towards a world that better reflects our values.

Given that work on fairness, ethics, and bias in AI is now part of the 'normative construction of the world' (Green & Hu, 2018, p. 5), sociologists can contribute to this world-building (Joyce et al., 2021) and help articulate possible directions for change that go beyond nebulous notions of an unbiased society or undesirable tendencies. Where vague gestures towards a 'social good' have predominated in AI ethics, there is an opportunity to specify desirable futures based on substantive equality and anti-oppression (Green, 2019; also; Davis et al., 2021). Unfortunately, while sociologists are on comfortable ground when critiquing representational inequalities and algorithmic harms, we are less comfortable in making normative proposals for what to do with algorithmic systems, beyond rejecting their use in problematic cases. A notable exception is Benjamin's (2019) *Race After Technology*, which does conclude with sections on 'abolitionist tools' and 'reimagining technology', but the book remains largely a work of explanation and critique and is very skeptical of 'design thinking'. Benjamin reminds us that when we are dealing with inequalities involving new technologies, we do not necessarily need a technological fix. Established forms of politics are often more appropriate, as expressed in the argument that 'maybe what we must demand is not liberatory *designs* but just plain old liberation' (p. 179). Keeping this in mind, what would it mean to enact systemic reforms that address inequality, in ways that go beyond rejection, refusal or opposition to sociotechnical systems? In other words, what positive actions are possible (involving both technology and politics) to create the sort of world we want?

While many students are attracted to sociology because of a desire for social change, those of us employed in academia often have other priorities, as reflected by the work we do and its audiences (Cancian, 1995; J. H. Turner, 2001; Weinstein, 2000). However, sociology does have a long history of arguing for social 'improvement' (Ward, 1906) and carrying out applied scholarship (Perlstadt, 2007). Along these lines, sociologists may engage in applied or participatory work with specific groups that are exercising agency through technology and data in order to meet their needs, or communities working towards some articulation of justice (Costanza-Chock, 2020). There are many groups whose interests have been marginalized in the design of new technologies, and their struggles are not as straight-forward as a fight for equality – encompassing goals that are specific to those communities. For example, in the context of settler colonialism (see Simpson, 2014), treating everyone on equal terms amounts to a form of assimilation for Indigenous peoples. The movement towards decolonization and Indigenous data sovereignty (Kukutai & Taylor, 2016; Lewis et al., 2020; Walter et al., 2020) is therefore not based on liberal concepts of individual equality, but the distinctive position of Indigenous peoples in regard to land rights, sovereignty, and cultural integrity.

Although there is a great deal of work that can be done to address the needs and desires of specific populations in relation to technologies, the fact that algorithmic systems are now implicated in inequalities 'at scale' also presents possibilities for more ambitious interventions. A national-level tax regime is one example of how a low-tech algorithm can broadly shift lines of inequality, but there are many more specific regimes for the distribution of wealth and resources that are in the process of being automated or 'transformed' through digital government (Clarke, 2020; Henman, 2010; Levy et al., 2021). Additionally, there are the algorithmic systems operated by massive platform companies (sometimes employing sociologists) that act as *de facto* governments, shaping outcomes for millions or billions of people (Gillespie, 2018; Tusikov, 2017). Sociologists can certainly seek to contribute to these algorithmic endeavors, ideally in an interdisciplinary manner, and mindful of the history of the relationship between sociology and technocracy, or social engineering. While partnering with established centres of power brings greater capacities to bear on the problems of algorithmic governance and inequality, the risk is that such work becomes subordinated to dominant interests and may further entrench the status quo.

5.2.1 | Interdisciplinary social engineering

Social engineering is a term that has typically been applied to top-down and expert-led projects to order society, particularly through state agencies. It can be exemplified by the policies of totalitarian governments of the twentieth century, including Nazi and communist regimes, but also programs in more socially democratic as well as liberal states, such as the eugenic policies that were once widespread in Europe and North America (Bashford & Levine, 2010; Lucassen, 2010). The relationship between sociology and social planning was a significant theme in the early years of the discipline (Ward, 1906), and while enthusiasm for social engineering faded in American social science over the 1930s (Jordan, 1994), it was addressed in the later work of Karl Mannheim (1940), as well as the relatively obscure discipline of 'sociotechnics', associated with Polish sociologist Adam Podgórecki (Podgórecki et al., 1996).

Today, few in North America would characterize themselves as 'social engineers', but in Sweden social engineering describes the establishment of the post-war welfare state. Although in Sweden the approach also largely fell out of favor later in the twentieth century, it remains in use to refer to 'the idea that we can create the greatest degree of happiness, the 'good society,' through rational social planning' (Etzemüller, 2014, p. 7). The development of social engineering has much in common with the birth of technocracy (Jordan, 1994), which also originated as an engineering approach to social problems, and uses specialized expertise or rational planning as a basis for decision-making (Esmark, 2020). While technocracy is likewise much-critiqued as a form of governance (particularly for its anti-democratic orientation, see Fischer, 1990) and there are few who would self-identify as technocrats, a technocratic rationality characterizes how many contemporary forms of governance operate across a variety of political systems (Esmark, 2017, 2020). Meanwhile, it has been claimed that through the potential of algorithmic technologies to reconfigure societies, 'software engineers will increasingly be the social engineers of the digital lifeworld' (Susskind, 2018, p. 294).

The future relationship between sociology and social engineering has been discussed by Jonathan Turner (2001, p. 101), who argued that sociology should develop an 'engineering wing' based on theoretical principles applied to real-world problems, and an ability to translate knowledge to 'fellow engineers' and 'clients'. Duncan J. Watts (2017) made a related argument as a sociologist employed at Microsoft, and this is also the kind of approach favored by a number of AI practitioners, imagining social scientists as problem solvers who can 'propose, implement, and evaluate' algorithmic processes to advance the 'social good' (Lepri et al., 2018). Data scientists and AI practitioners are more likely to look to social science for solutions to an existing problem, than to consider how problems or perspectives in social science translate to data science. For example, data scientists may have questions about human behaviour or cognition that they imagine social scientists can design experiments to answer (Irving & Askell, 2019). Collaborations between data science and social science around a specific topic or shared problem often end up being led by technically-skilled participants, with programmers deciding what counts as important, while less-technical participants take a back seat. This produces familiar 'disciplinary divisions of labor: social scientists observe, data scientists make; social scientists do ethics, data scientists do science' (Moats & Seaver, 2019, p. 8). In these processes, disciplinary boundaries are maintained and social science plays a supporting or 'additive' (Lury, 2018, p. 1) role in contributing knowledge.

In contrast to *multidisciplinary* work as described above, *interdisciplinarity* can be considered more of a reciprocal, transformative, or integrative relationship between disciplines. According to Lury, interdisciplinary methods are 'dynamic conduits for relations of interference in which differences and asymmetries between disciplines are explored and exploited in relation to specific problems, in specific places, with specific materials' (2018, p. 21). Here, the emphasis is on crossing disciplines to 'constitute' problems in new ways, rather than arriving at novel solutions, and interdisciplinarity is less an instrumental practice and more of an enterprise in research autonomy (Lury, 2018). Even in situations where two disciplines use different language to discuss what might appear to be the same problem, switching from one disciplinary discourse to another can significantly shift how problems are formulated. This is certainly the case when it comes to questions of bias and fairness in AI, once we attempt to translate these into discourses of social inequality. 'Reducing bias' is not at all synonymous with 'reducing oppression'. Researchers can attempt to correct for 'societal bias' (Friedman & Nissenbaum, 1996; Silberg & Manyika, 2019), but understanding the source of this structural inequality is typically treated as irrelevant, or outside the scope of analysis. Social structures, or 'sources of discrimination that cannot be traced to discrete bad mechanisms are bracketed, dismissed as someone else's problem or, worse, couched as untouchable facts of history' (Hoffmann, 2019, p. 905). Sociologists can play a valuable role in the formulation of problems as well as in identifying where interventions may be effective, but interdisciplinary openness will be required from collaborators to make this possible.

For AI researchers and data scientists, the key problem in recent years has been eliminating unfair discrimination and producing 'fair' results through automated decisions. If sociologists are brought in to help once a sociotechnical system has already been proposed and the problem of fair decision-making raised, the scope of our agency is likely to be quite limited. Sociologists may have more experience with philosophy and politics than those working in technical fields, but this does not confer an ability to configure a 'fair' distribution from an established decision-making system. Sociology's strengths include our theoretical and empirical analyses of inequality – an understanding of systems of stratification and forms of discrimination. Systemic problems such as racism and poverty are structurally reproduced and require systemic approaches; sociologists can help to articulate what these approaches might look like (e.g., Davis et al.'s [2021] proposal for 'algorithmic reparation'), or where to focus our energies to achieve social change. Therefore, the greatest contributions that sociologists can make are at earlier stages in the development of policies and sociotechnical systems, where goals and techniques are less clearly defined. Once plans are underway, sociologists are likely to be enrolled in ways that are complicit with organizational goals such as efficiency, legitimacy, or profit (Maris, 2022).

5.3 | Governance of algorithmic governance

Issues related to social inequality are of collective or public interest, and therefore are often pursued through institutions responsible for the public good. This leads to the third way that sociologists can be involved in the normative construction of this world, by participating in the governance of technology, and addressing the very real dangers of algorithmic governance.

Recent years have seen a considerable amount of scholarship documenting, comparing and critiquing different policies or regimes for regulating algorithms and AI, including government strategies, corporate statements of principles, standards, and regulations (Bradley et al., 2020; Jobin et al., 2019; Stark et al., 2021; J. Turner, 2019). A positive approach builds on this work to develop and improve these governance regimes. Seyfert (2021) argues that regulatory processes co-produce the algorithms they regulate, but AI-specific regulations are a new development, often lacking 'teeth' or meaningful enforcement. To the extent that formal political channels are open to us, sociologists can contribute to the development of the nascent regulatory regimes being established to govern AI and algorithms in their respective jurisdictions or in an international forum.

The degree to which sociologists can contribute to public policy on these issues depends on the political opportunities available, such as whether formal processes are open to academic input, or the extent to which consultations are actually used to inform policy (rather than playing a legitimating or performative role, see Kerr et al., 2020). Critical experts who are excluded from the policy process can establish their own organizations, as a number of leading scholars did in 2020 with the 'Real Facebook Oversight Board' (Solon, 2020) – but these efforts remain at the level of policy critique rather than policy making, and can be ignored by industry. The current moment shows a strong appetite for government regulation of 'tech giants' in much of the world (Lee, 2021), and sociologists can contribute by emphasizing questions of power, political economy and structural inequalities in these debates. While regulators grapple with addressing the power that the corporate leaders in AI have achieved in recent years, the greatest societal 'disruptions' will likely come through the use of these technologies by government agencies (CSPS, 2021). Although the algorithms deployed by Google, Amazon, or Facebook affect billions of people, algorithms used in the public sector can have more profound consequences – and are also more open to contestation and study by independent researchers.

Sociologists need to study these emerging uses of AI in public institutions, while also fighting to keep them publicly accountable. I have been documenting the use of ADM systems in Canada, where government agencies have been employing these technologies for hiring decisions, benefits claims, immigration applications, legal analysis, sentiment analysis, suicide prediction, fraud detection, facial recognition, and public-serving chatbots (among many other uses – see ESDC, 2019; PWGSC, 2018; Reevely, 2021). Governing such algorithmic systems may include the creation of new regulatory processes and agencies, but AI and algorithms are already regulated through privacy and data protection law, competition law, human rights and anti-discrimination law, and more specialized domains where these technologies are applied, such as medicine and public administration. Currently, algorithmic and AI regulations have more to do with questions of transparency, accountability, privacy, competition, and fairness or justice, than social inequality as a policy problem, but this could be changed by orienting current issues in AI regulation toward the pursuit of substantive equality (Amani, 2021). Critical and engaged scholarship is needed both at the level of broad regulatory regimes, as well as the specific implementation of these sociotechnical systems in particular domains, which may have their own specific channels for democratic accountability and public participation.

Algorithmic governance therefore presents sociologists with a research problem, and a way to make a positive contribution to policy. Academic involvement can include carefully documenting algorithmic processes, how they affect human lives and outcomes, critiquing and opposing the harms they cause, and reforming, improving, or constraining these systems through regulatory and public policy interventions. Broadly speaking, sociologists can help define the role of automated systems in a democratic society. In comparison to twentieth-century social engineering projects, today's technocratic 'innovations' are more frequently presented as compatible with (and even bolstering) democratic processes (Esmark, 2020), but the historical critiques of technocracy in a democratic context (Jordan, 1994) remain relevant, including the reliance on experts to determine the public interest and the dangers of an ideological pursuit of efficiency. The use of automated systems to implement public policy in a democracy raises major issues that require closer attention from all citizens, sociologists included.

6 | CONCLUSION

Sociologists must now grapple with the role of algorithmic systems and AI in the distribution of various goods and outcomes; these systems are widespread today, and will continue to become even more central to the reproduction of inequality in the future. Given the enormity of the ethical and political challenges involved, and the limitations that other fields have demonstrated in addressing them, sociologists have an opportunity to make a valuable contribution to human welfare. Specifically, issues that are currently presented as problems with bias can often be understood through social theory as problems of inequality, made manifest in structures that reproduce benefits for some and harms or cumulative disadvantages for others.

Sociologically-informed critiques of technologies have already had some influence on current debates but remain in need of further elaboration. When critique extends into the politics of refusal, alliances can be formed with others who are motivated to reject specific developments and promote more positive ideas for social change, including alternative sociotechnical systems that are compatible with our visions and values. This might mean working with groups and communities that are building technologies to meet their needs from the bottom-up, but it can also mean critically engaging with systems explicitly intended to serve the public good – namely through the public sector. Relatedly, broader democratic and interdisciplinary participation is needed in the development of new public policies and governance mechanisms that will help co-produce the algorithms of the future.

This will require a practical orientation from sociologists, working beyond our discipline, and articulating normative positions and policy proposals that others will hear. Thankfully, much of the heavy lifting has already been done by critical voices behind the 'techlash' of recent years. Algorithmic harms, racist and sexist robots, ubiquitous surveillance and behavioral manipulation have now all become well-established as public issues. Artificial intelligence researchers sometimes assume that social scientists can help them address these issues, and sociologists can indeed choose to play such a collaborative role, as long as we are mindful of the limits of our agency and the danger of our contributions becoming subordinated to other interests. Active engagement with algorithmic systems does not necessitate sociologists who can write code, but it does force us to consider the discipline's relationship to governance and politics, including where we have the greatest possibilities of effecting change. Just as sociological critique can be effective by locating technologies and individualized consequences in systemic terms or a larger context, applied sociology may help orient these sociotechnical systems to address larger, interrelated, and systemic problems, including various kinds of inequality and their consequences.

ACKNOWLEDGMENT

None.

ORCID

Mike Zajko https://orcid.org/0000-0001-7804-4618

REFERENCES

Airoldi, M. (2022). *Machine habitus: Toward a sociology of algorithms*. Polity Press.

Amani, B. (2021). AI and equality by design. In F. Martin-Bariteau & T. Scassa (Eds.), *Artificial intelligence and the law in Canada* (pp. 267–300).

Anderson, J., & Rainie, L. (2018). *Artificial intelligence and the future of humans*. Pew Research Center. https://www.pewresearch.org/internet/2018/12/10/artificial-intelligence-and-the-future-of-humans/

Aneesh, A. (2009). Global labor: Algocratic modes of organization. *Sociological Theory*, *27*(4), 347–370. https://doi.org/10.1111/j.1467-9558.2009.01352.x

Bailey, S., Pierides, D., Brisley, A., Weisshaar, C., & Blakeman, T. (2020). Dismembering organisation: The coordination of algorithmic work in healthcare. *Current Sociology*, *68*(4), 546–571. https://doi.org/10.1177/0011392120907638

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, *20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Bashford, A. & Levine, P. (Eds.). (2010). *The Oxford handbook of the history of eugenics*. Oxford University Press.

Bastani, A. (2019). *Fully automated luxury communism: A manifesto*. Verso Books.

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

Bhambra, G. K. (2014). Knowledge production in global context: Power and coloniality [Special issue]. *Current Sociology*, *62*(4).

Birhane, A. (2020). *Fair warning. Real life*. https://reallifemag.com/fair-warning/

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2), 1–9. https://doi.org/10.1016/j.patter.2021.100205

Boyd, R., & Holton, R. J. (2018). Technology, innovation, employment and power: Does robotics and artificial intelligence really mean social transformation? *Journal of Sociology*, *54*(3), 331–345. https://doi.org/10.1177/1440783317726591

Bradley, C., Wingfield, R., & Metzger, M. (2020). *National artificial intelligence strategies and human rights: A review*. https://www.gp-digital.org/wp-content/uploads/2020/04/National-Artifical-Intelligence-Strategies-and-Human-Rights%E2%80%94A-Review_April2020.pdf

Brayne, S. (2020). *Predict and surveil: Data, discretion, and the future of policing*. Oxford University Press.

Brayne, S., & Christin, A. (2021). Technologies of crime prediction: The reception of algorithms in policing and criminal courts. *Social Problems*, *68*, 608–624.

Burrell, J., & Fourcade, M. (2021). The society of algorithms. *Annual Review of Sociology*, *47*(1), 213–237. https://doi.org/10.1146/annurev-soc-090820-020800

Cancian, F. M. (1995). Truth and goodness: Does the sociology of inequality promote social betterment? *Sociological Perspectives*, *38*(3), 339–356. https://doi.org/10.2307/1389431

Christin, A. (2020). The ethnographer and the algorithm: Beyond the black box. *Theory and Society*, *49*(5), 897–918.

CIFAR. (2020). *DLRL summer school 2020—Ethics in AI - panel discussion*. https://www.youtube.com/watch?v=ThYslOokwR4

Cifor, M., Garcia, P., Cowan, T. L., Rault, J., Sutherland, T., Chan, A. S., Rode, J., Hoffmann, A. L., Salehi, N., & Nakamura, L. (2019). *Feminist data manifest-no*. https://www.manifestno.com/home

Clarke, A. (2020). Digital government units: What are they, and what do they mean for digital era public management renewal? *International Public Management Journal*, *23*(3), 358–379. https://doi.org/10.1080/10967494.2019.1686447

Clifford, C. (2021). *OpenAI's Sam Altman: Artificial Intelligence will generate enough wealth to pay each adult $13,500 a year*. CNBC. https://www.cnbc.com/2021/03/17/openais-altman-ai-will-make-wealth-to-pay-all-adults-13500-a-year.html

Collins, H. (2018). *Artifictional intelligence: Against humanity's surrender to computers*. Polity Press.

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

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

Crawford, K. (2021). *Atlas of AI*. Yale University Press.

Cruz, T. M. (2020). Perils of data-driven equity: Safety-net care and big data's elusive grasp on health inequality. *Big Data & Society*, *7*(1), 1–14. https://doi.org/10.1177/2053951720928097

CSPS (2021). *CSPS virtual café series: Artificial intelligence*. https://www.youtube.com/watch?v=6drggKeFfC4

Davies, H. C., Eynon, R., & Salveson, C. (2021). The mobilisation of AI in education: A bourdieusean field analysis. *Sociology*, *55*(3), 539–560. https://doi.org/10.1177/0038038520967888

Davis, J. L., Williams, A., & Yang, M. W. (2021). Algorithmic reparation. *Big Data & Society*, *8*(2), 1–14. https://doi.org/10.1177/20539517211044808

Denton, E., Hanna, A., Amironesei, R., Smart, A., & Nicole, H. (2021). On the genealogy of machine learning datasets: A critical history of ImageNet. *Big Data & Society*, *8*(2), 1–14. https://doi.org/10.1177/20539517211035955

DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). Digital inequality: From unequal access to differentiated use. In K. M. Neckerman (Ed.), *Social inequality* (pp. 355–400). Russell Sage Foundation. https://doi.org/10.7758/9781610444200.14

Dyer-Witheford, N., Kjøsen, A. M., & Steinhoff, J. (2019). *Inhuman power: Artificial intelligence and the future of capitalism*. Pluto Press.

Elliott, A. (2019). *The culture of AI: Everyday life and the digital revolution*. Routledge.
Elliott, A. (Ed.). (2022). *The Routledge social science handbook of AI*. Routledge.
Erigha, M. (2015). Race, gender, Hollywood: Representation in cultural production and digital media's potential for change. *Sociology Compass*, *9*(1), 78–89. https://doi.org/10.1111/soc4.12237
ESDC. (2019). *Final or most recent version of Employment and Social Development Canada artificial intelligence strategy (access to information request A-2019-01027)*.
Esmark, A. (2017). Maybe it is time to rediscover technocracy? An old framework for a new analysis of administrative reforms in the governance era. *Journal of Public Administration Research and Theory*, *27*(3), 501–516. https://doi.org/10.1093/jopart/muw059
Esmark, A. (2020). *The new technocracy*. Bristol University Press.
Etzemüller, T. (2014). *Alva and Gunnar Myrdal: Social engineering in the modern world*. Lexington Books.
Eubanks, V. (2017). *Automating inequality: How high-tech tools profile, police, and punish the poor*. Martin's Press.
Felt, U., Fouché, R, Miller, C. A., & Smith-Doerr, L. (Eds.). (2017). *The handbook of science and technology studies* (4th ed.). MIT Press.
Fischer, F. (1990). *Technocracy and the politics of expertise*. Sage.
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
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
Gibbs, S. (2016). *Google alters search autocomplete to remove "are Jews evil" suggestion*. The Guardian. https://www.theguardian.com/technology/2016/dec/05/google-alters-search-autocomplete-remove-are-jews-evil-suggestion
Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
Gillespie, T., & Seaver, N. (2016). *Critical algorithm studies: A reading list. Social media collective*. http://socialmediacollective.org/reading-lists/critical-algorithm-studies/
Go, J. (2020). Race, empire, and epistemic exclusion: Or the structures of sociological thought. *Sociological Theory*, *38*(2), 79–100. https://doi.org/10.1177/0735275120926213
Grant, N., Bass, D., & Eidelson, J. (2021). *Google Ethical AI group's turmoil began long before public unraveling*. Bloomberg. https://www.bloomberg.com/news/articles/2021-04-21/google-ethical-ai-group-s-turmoil-began-long-before-public-unraveling
Gray, M. L., & Suri, S. (2019). *Ghost work: How to stop Silicon Valley from building a new global underclass*. Houghton Mifflin Harcourt.
Green, B. (2018). *Data science as political action: Grounding data science in a politics of justice*. ArXiv Preprint ArXiv:1811.03435. https://arxiv.org/abs/1811.03435
Green, B. (2019). "Good" isn't good enough. *Proceedings of the AI for social good workshop at NeurIPS*. https://www.benzevgreen.com/wp-content/uploads/2019/11/19-ai4sg.pdf
Green, B., & Hu, L. (2018). *The myth in the methodology: Towards a recontextualization of fairness in machine learning. 35th international conference on machine learning*. https://scholar.harvard.edu/files/bgreen/files/18-icmldebates.pdf
Griesbach, K., Reich, A., Elliott-Negri, L., & Milkman, R. (2019). Algorithmic control in platform food delivery work. *Socius*, *5*, 2378023119870041. https://doi.org/10.1177/2378023119870041
Grohmann, R., & Araújo, W. F. (2021). Beyond Mechanical Turk: The work of Brazilians on global AI platforms. In P. Verdegem (Ed.), *AI for everyone? Critical perspectives* (pp. 247–266). University of Westminster Press.
Hacker, P. (2018). Teaching fairness to artificial intelligence: Existing and novel strategies against algorithmic discrimination under EU law. *Common Market Law Review*, *55*(4), 1143–1185.
Hanna, A., Denton, E., Smart, A., & Smith-Loud, J. (2019). *Towards a critical race methodology in algorithmic fairness. Conference on Fairness, Accountability, and Transparency (FAT* '20)*. https://doi.org/10.1145/3351095.3372826
Hashemi, M., & Hall, M. (2020). Criminal tendency detection from facial images and the gender bias effect. *Journal of Big Data*, *7*(2), 1–16. https://doi.org/10.1186/s40537-019-0282-4
Helsper, E. J. (2008). *Digital inclusion: An analysis of social disadvantage and the information society*. Department for Communities and Local Government. http://www.communities.gov.uk/documents/communities/pdf/digitalinclusionanalysis
Helsper, E. J. (2012). A corresponding fields model for the links between social and digital exclusion. *Communication Theory*, *22*(4), 403–426. https://doi.org/10.1111/j.1468-2885.2012.01416.x
Henman, P. (2010). *Governing electronically e-government and the reconfiguration of public administration, policy and power*. Palgrave Macmillan.
Hoffman, S. G. (2021). A story of nimble knowledge production in an era of academic capitalism. *Theory and Society*, *50*(4), 541–575. https://doi.org/10.1007/s11186-020-09422-0
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

Iliadis, A., & Russo, F. (2016). Critical data studies: An introduction. *Big Data & Society*, *3*(2), 1–7. https://doi.org/10.1177/2053951716674238

Irani, L., & Chowdhury, R. (2019). *To really "disrupt," tech needs to listen to actual researchers*. Wired. https://www.wired.com/story/tech-needs-to-listen-to-actual-researchers/

Irving, G., & Askell, A. (2019). AI safety needs social scientists. *Distill*, *4*. https://doi.org/10.23915/distill.00014

Issar, S., & Aneesh, A. (2022). What is algorithmic governance? *Sociology Compass*, *16*(1), e12955. https://doi.org/10.1111/soc4.12955

James, A., & Whelan, A. (2021). 'Ethical' artificial intelligence in the welfare state: Discourse and discrepancy in Australian social services. *Critical Social Policy*, *42*, 1–42. https://doi.org/10.1177/0261018320985463

Jaton, F. (2021). *The constitution of algorithms: Ground-truthing, programming, formulating*. MIT Press.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jordan, J. M. (1994). *Machine-age ideology: Social engineering and American liberalism, 1911-1939*. University of North Carolina Press.

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a sociology of artificial intelligence: A call for research on inequalities and structural change. *Socius*, *7*, 1–11. https://doi.org/10.1177/2378023121999581

Kalluri, P. (2020). Don't ask if artificial intelligence is good or fair, ask how it shifts power. *Nature*, *583*(7815), 169. https://doi.org/10.1038/d41586-020-02003-2

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *The Academy of Management Annals*, *14*(1), 366–410.

Kerr, A., Barry, M., & Kelleher, J. D. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. *Big Data & Society*, *7*(1), 1–12. https://doi.org/10.1177/2053951720915939

Kind, C. (2020). *The term 'ethical AI' is finally starting to mean something*. VentureBeat. https://venturebeat.com/2020/08/23/the-term-ethical-ai-is-finally-starting-to-mean-something/

Kukutai, T. & Taylor, J. (Eds.). (2016). *Indigenous data sovereignty*. ANU Press. https://doi.org/10.22459/CAEPR38.11.2016

Kusner, M. J., & Loftus, J. R. (2020). The long road to fairer algorithms. *Nature*, *578*(7793), 34–36. https://doi.org/10.1038/d41586-020-00274-3

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

Lee, T. B. (2021). *Why Big Tech is facing regulatory threats from Australia to Arizona*. Ars Technica. https://arstechnica.com/tech-policy/2021/03/why-big-tech-is-facing-regulatory-threats-from-australia-to-arizona/

Lempert, R. (2019). Bezos World or Levelers: Can we choose our scenario? https://www.rand.org/pubs/external_publications/EP67822.html

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, *31*(4), 611–627. https://doi.org/10.1007/s13347-017-0279-x

Levy, K., Chasalow, K., & Riley, S. (2021). Algorithms and decision-making in the public sector. *Annual Review of Law and Social Science*, *17*, 309–334. https://doi.org/10.1146/annurev-lawsocsci-041221-023808

Lewis, J. E., Abdilla, A., Arista, N., Baker, K., Benesiinaabandan, S., Brown, M., Cheung, M., Coleman, M., Cordes, A., Davison, J., Duncan, K., Garzon, S., Harrell, D. F., Jones, P.-L., Kealiikanakaoleohaililani, K., Kelleher, M., Kite, S., Lagon, O., Leigh, J., … Whaanga, H. (2020). *Indigenous protocol and artificial intelligence position paper*. Indigenous protocol and artificial intelligence working group and the Canadian institute for advanced research. https://spectrum.library.concordia.ca/986506/

Lucassen, L. (2010). A brave new world: The left, social engineering, and eugenics in twentieth-century Europe. *International Review of Social History*, *55*(2), 265–296. https://doi.org/10.1017/S0020859010000209

Lury, C. (2018). Introduction: Activating the present of interdisciplinary methods. In C. Lury, R. Fensham, A. Heller-Nicholas, S. Lammes, A. Last, M. Michael & E. Uprichard (Eds.), *Routledge handbook of interdisciplinary research methods* (pp. 1–25). Routledge. https://doi.org/10.4324/9781315714523

Lyon, D. (2003). In *Surveillance as social sorting: Privacy, risk, and digital discrimination*. Routledge.

Mannheim, K. (1940). *Man and society in an age of reconstruction: Studies in modern social structure*. Kegan Paul. (E. Shils, Trans.).

Maris, E. (2022). *The humanities can't save big tech from itself. Wired*. https://www.wired.com/story/ethicis-big-tech-humanities/

McGranahan, C. (2016). Theorizing refusal: An introduction. *Cultural Anthropology*, *31*(3), 319–325.

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). *A survey on bias and fairness in machine learning*. ArXiv:1908.09635 [Cs]. http://arxiv.org/abs/1908.09635

Milan, S., & Treré, E. (2019). Big data from the south(s): Beyond data universalism. *Television & New Media*, *20*(4), 319–335. https://doi.org/10.1177/1527476419837739

Mitchell, S., Potash, E., Barocas, S., D'Amour, A., & Lum, K. (2020). *Prediction-based decisions and fairness: A catalogue of choices, assumptions, and definitions*. ArXiv:1811.07867 [Stat]. http://arxiv.org/abs/1811.07867

Moats, D., & Seaver, N. (2019). “You social scientists love mind games”: Experimenting in the “divide” between data science and critical algorithm studies. *Big Data & Society*, *6*(1), 1–11. https://doi.org/10.1177/2053951719833404
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
Newlands, G. (2021). Lifting the curtain: Strategic visibility of human labour in AI-as-a-Service. *Big Data & Society*, *8*(1), 1–14. https://doi.org/10.1177/20539517211016026
Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
Ossewaarde, M., & Gulenc, E. (2020). National varieties of artificial intelligence discourses: Myth, utopianism, and solutionism in West European policy expectations. *Computer*, *53*(11), 53–61. https://doi.org/10.1109/MC.2020.2992290
Park, S., & Humphry, J. (2019). Exclusion by design: Intersections of social, digital and data exclusion. *Information, Communication & Society*, *22*(7), 934–953. https://doi.org/10.1080/1369118X.2019.1606266
Pascale, C.-M. (Ed.). (2013). *Social inequality & the politics of representation*. SAGE.
Perlstadt, H. (2007). Applied sociology. In C. D. Bryant & D. L. Peck (Eds.), *21st century sociology* (pp. 342–352). SAGE.
Podgórecki, A., Alexander, J., & Shields, R. (Eds.). (1996). *Social engineering*. Carleton University Press.
PWGSC. (2018). *Artificial Intelligence (AI) pilot project for surveillance of suicide-related behaviours using social media* (1000203413). https://buyandsell.gc.ca/procurement-data/tender-notice/PW-18-00840702
Reevely, D. (2021). *Federal rules on AI too narrow and risk ‘damaging public trust’: Internal review. The Logic*. https://thelogic.co/news/federal-rules-on-ai-too-narrow-and-risk-damaging-public-trust-internal-review/
Roberge, J., & Castelle, M. (2021). Toward an end-to-end sociology of 21st-century machine learning. In J. Roberge & M. Castelle (Eds.), *The cultural life of machine learning: An incursion into critical AI studies* (pp. 1–29). Springer International Publishing.
Roberts, D. E. (2019). Digitizing the carceral state. *Harvard Law Review*, *132*(6), 1695–1728.
Roberts, D. E., & Rollins, O. (2020). Why sociology matters to race and biosocial science. *Annual Review of Sociology*, *46*, 195–214.
Sampath, P. G. (2021). Governing artificial intelligence in an age of inequality. *Global Policy*, *12*(S6), 21–31. https://doi.org/10.1111/1758-5899.12940
Scheuerman, M. K., Pape, M., & Hanna, A. (2021). Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society*, *8*(2), 1–15. https://doi.org/10.1177/20539517211053712
Schou, J., & Pors, A. S. (2019). Digital by default? A qualitative study of exclusion in digitalised welfare. *Social Policy and Administration*, *53*(3), 464–477. https://doi.org/10.1111/spol.12470
Schuilenburg, M. & Peeters, R. (Eds.). (2021). *The algorithmic society: Technology, power, and knowledge*. Routledge.
Science Academies, G. 7 (2019). *Artificial intelligence and society*. https://rscsrc.ca/sites/default/files/Artificial%20intelligence%20and%20society%20G7%202019.pdf
Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. *New Media & Society*, *6*(3), 341–362. https://doi.org/10.1177/1461444804042519
Seyfert, R. (2021). Algorithms as regulatory objects. *Information, Communication & Society*, *0*(0), 1–17. https://doi.org/10.1080/1369118X.2021.1874035
Shah, A. (2018). *AI for the good of society*. BBC Blogs. https://www.bbc.co.uk/blogs/internet/entries/7e49f841-85af-4455-a8b0-c16e6279176c
Shestakofsky, B. (2017). Working algorithms: Software automation and the future of work. *Work and Occupations*, *44*(4), 376–423. https://doi.org/10.1177/0730888417726119
Shestakofsky, B. (2020). Stepping back to move forward: Centering capital in discussions of technology and the future of work. *Communication and the Public*, *5*(3–4), 129–133. https://doi.org/10.1177/2057047320959854
Shestakofsky, B., & Kelkar, S. (2020). Making platforms work: Relationship labor and the management of publics. *Theory and Society*, *49*(5), 863–896. https://doi.org/10.1007/s11186-020-09407-z
Silberg, J., & Manyika, J. (2019). *Notes from the AI frontier: Tackling bias in AI (and in humans)*. https://www.mckinsey.com/~/media/mckinsey/featured insights/artificial intelligence/tackling bias in artificial intelligence and in humans/mgi-tackling-bias-in-ai-june-2019.ashx
Simonite, T. (2021). *What really happened when google ousted Timnit Gebru. Wired*. https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened
Simpson, A. (2014). *Mohawk interruptus: Political life across the borders of settler states*. Duke University Press.
Sinanan, J., & McNamara, T. (2021). Great AI divides? Automated decision-making technologies and dreams of development. *Continuum*, *35*, 1–760. https://doi.org/10.1080/10304312.2021.1983257
Singh, S., & Steeves, V. (2020). The contested meanings of race and ethnicity in medical research: A case study of the DynaMed point of care tool. *Social Science & Medicine*, *265*, 113112. https://doi.org/10.1016/j.socscimed.2020.113112
Solon, O. (2020). *While Facebook works to create an oversight board, industry experts formed their own*. NBC News. https://www.nbcnews.com/tech/tech-news/facebook-real-oversight-board-n1240958

Stark, L., Greene, D., & Hoffmann, A. L. (2021). Critical perspectives on governance mechanisms for AI/ML systems. In J. Roberge & M. Castelle (Eds.), *The cultural life of machine learning: An incursion into critical AI studies* (pp. 257–280). Springer International Publishing.
Stark, L., & Hutson, J. (forthcoming). Physiognomic artificial intelligence. *Fordham Intellectual Property, Media and Entertainment Law Journal*. https://doi.org/10.2139/ssrn.3927300
Stinson, C. (2021). The dark past of algorithms that associate appearance and criminality. *American Scientist*, *109*(1), 26–29.
Suresh, H., & Guttag, J. V. (2020). *A framework for understanding unintended consequences of machine learning*. ArXiv:1901.10002 [Cs, Stat]. http://arxiv.org/abs/1901.10002
Susskind, J. (2018). *Future politics: Living together in a world transformed by tech*. Oxford University Press.
Turner, J. (2019). *Robot rules: Regulating artificial intelligence*. Palgrave Macmillan.
Turner, J. H. (2001). Social engineering: Is this really as bad as it sounds? *Sociological Practice*, *3*(2), 99–120.
Tusikov, N. (2017). *Chokepoints: Global private regulation on the internet*. University of California Press.
van Dijk, J. A. G. M. (2005). *The deepening divide: Inequality in the information society*. SAGE.
van Dijk, J. A. G. M. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4–5), 221–235. https://doi.org/10.1016/j.poetic.2006.05.004
Vicsek, L. (2020). Artificial intelligence and the future of work – lessons from the sociology of expectations. *International Journal of Sociology & Social Policy*, *41*(7/8), 842–861. https://doi.org/10.1108/IJSSP-05-2020-0174
Wajcman, J. (2017). Automation: Is it really different this time? *British Journal of Sociology*, *68*(1), 119–127. https://doi.org/10.1111/1468-4446.12239
Walch, K. (2020). *Why the race for AI dominance is more global than you think*. Forbes. https://www.forbes.com/sites/cognitiveworld/2020/02/09/why-the-race-for-ai-dominance-is-more-global-than-you-think/
Walter, M., Kukutai, T., Carroll, S. R., & Rodriguez-Lonebear, D. (Eds.). (2020). *Indigenous data sovereignty and policy*. Routledge. https://doi.org/10.4324/9780429273957
Ward, L. F. (1906). *Applied sociology: A treatise on the conscious improvement of society by society*. Ginn & Company.
Watts, D. J. (2017). Should social science be more solution-oriented? *Nature Human Behaviour*, *1*(1), 1–5. https://doi.org/10.1038/s41562-016-0015
Weinstein, J. (2000). The place of theory in applied sociology: A reflection. *Theory & Science*, *1*(1). https://theoryandscience.icaap.org/content/vol001.001/01weinstein_revised.html
Wellman, B. & Haythornthwaite, C. (Eds.). (2002). *The internet in everyday life*. Blackwell.
West, K. G., Schechner, S., McMillan, R., John. (2019). How Google interferes with its search algorithms and changes your results. *Wall Street Journal*. https://www.wsj.com/articles/how-google-interferes-with-its-search-algorithms-and-changes-your-results-11573823753
Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995
Winner, L. (1980). Do artifacts have politics? *Dædalus*, *109*(1), 121–136.
Zajko, M. (2021). Conservative AI and social inequality: Conceptualizing alternatives to bias through social theory. *AI & Society*, *36*(3), 1047–1056. https://doi.org/10.1007/s00146-021-01153-9
Zarkadakis, G. (2020). *Cyber republic: Reinventing democracy in the age of intelligent machines*. MIT Press.
Zuboff, S. (2018). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

AUTHOR BIOGRAPHY

Mike Zajko is an Assistant Professor in the Department of History and Sociology, University of British Columbia, Okanagan Campus. He holds a PhD in Sociology from the University of Alberta. His current research examines automated decision-making in the public sector and the use of machine-readable rules to govern behaviour.

How to cite this article: Zajko, M. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass*, *16*(3), e12962. https://doi.org/10.1111/soc4.12962