

Are Ethics Enough in Data Science?

This chapter draws on material from:

- [*Data Science as Political Action: Grounding Data Science in a Politics of Justice*](#) by [Ben Green](#), licensed under [CC BY 4.0](#)
- [2. Collect, Analyze, Imagine, Teach](#) by [Catherine D'Ignazio](#) and [Lauren Klein](#), licensed under [CC BY 4.0](#)

Changes to the source material include removal of original material, reformatting original material, addition of new material, combining of sources, and editing of original material for a different audience.

Introduction

The field of data science has entered a period of reflection and reevaluation. Alongside its rapid growth in both size and stature in recent years, data science has become beset by controversies and scrutiny. Machine learning algorithms that guide decisions in areas such as hiring, healthcare, criminal sentencing, and welfare are often biased, inscrutable, and proprietary (Anwin et al., 2016; Buolamwini & Gebru, 2018; Eubanks, 2018; Obermeyer et al., 2019; O'Neil, 2017; Wexler, 2018). Algorithms that drive social media feeds manipulate people's emotions (Kramer et al., 2014), spread misinformation (Vosoughi et al., 2018), and amplify political extremism (Nicas, 2018). Facilitating these and other algorithms are massive datasets, often gained illicitly or without meaningful consent, that reveal sensitive and intimate information about people (de Montjoye et al., 2015; Kosinski et al., 2013; Rosenberg et al., 2018; Thompson & Warzel, 2019)

Many individuals and organizations responded to these controversies by advocating for a focus on ethics in training and practice (Green, 2021). Data ethics represents a growing interdisciplinary effort—both critical and computational—to ensure that the ethical issues brought about by our increasing reliance on data-driven systems are identified and addressed. Thus far, the major trend has been to emphasize the issue of “bias,” and the values of “fairness, accountability, and transparency” in mitigating its effects. The broad motivation behind these efforts is the assumption that, if only data scientists were more attuned to the ethical implications of their work, many harms associated with data science could be avoided.

This is a promising development, especially for technical fields that have not historically foregrounded ethical issues, and as funding mechanisms for research on data and ethics proliferate. However, addressing bias in a dataset is a tiny technological Band-Aid for a much larger problem. Even the values of “fairness, accountability, and transparency,” which seek to address instances of bias in data-driven systems, are themselves non-neutral, as they locate the source of the bias in individual people and specific design decisions.

In short, these concepts may do good work, but they ultimately keep the roots of the problem in place. In other words, they maintain the current structure of power, even if they don't intend to, because they let structural issues off the hook. They direct data scientists' attention toward seeking technological fixes instead of social and political solutions. Sometimes those fixes are necessary and important. But as technology scholars Julia Powles (2018) and Helen Nissenbaum assert, “Bias is real, but it's also a captivating diversion.” There is a more fundamental problem that must

also be addressed: we do not all arrive in the present with equal power or privilege. Hundreds of years of history and politics and culture have brought us to the present moment. This is a reality of our lives as well as our data.

Ethics vs. Politics

In practice, technology ethics suffer from four significant limitations (Green, 2021). First, technology ethics principles are abstract and lack mechanisms to ensure that engineers follow ethical principles. Second, technology ethics has a myopic focus on individual engineers and on technology design, overlooking the structural sources of technological harms. Third, technology ethics is subsumed into corporate logics and practices rather than substantively altering behavior. All told, the rise of technology ethics often reflects a practice dubbed “ethics-washing”: tech companies deploying the language of ethics to resist more structural reforms that would curb their power and profits.

Thus, while ethics provides useful frameworks to help data scientists reflect on their practice and the impacts of their work, these approaches are insufficient for generating a data science that avoids social harms and that promotes social justice. The normative responsibilities of data scientists cannot be managed through a narrow professional ethics that lacks normative weight and supposes that, with some reflection and a commitment to best practices, data scientists will make the “right” decisions that lead to “good” technology. Instead of relying on vague moral principles that obscure the structural drivers of injustice, data scientists must engage in politics: the process of negotiating between competing perspectives, values, and goals.

In other words, we must recognize data science as a form of political action. Data scientists must recognize themselves as political actors engaged in normative constructions of society. In turn, data scientists must evaluate their efforts according to the downstream impacts on people’s lives.

In this context, *politics* and *political* do not refer to partisan or electoral debates about specific parties and candidates. Instead, these terms have a broader meaning that transcends activity directly pertaining to the government, its laws, and its representatives. Two aspects of politics are particularly important:

- First, politics is everywhere in the social world. As defined by politics professor Adrian Leftwich (1984), “politics is at the heart of all collective social activity, formal and informal, public and private, in all human groups, institutions, and societies”.
- Second, politics has a broad reach. Political scientist Harold Lasswell (1936) describes politics as “who gets what, when, how.” The “what” here could mean many things: money, goods, status, influence, respect, rights, and so on.

Understood in these terms, politics comprises any activities that affect or make claims about the who, what, when, and how in social groups, both small and large.

Data scientists are political actors in that they play an increasingly powerful role in determining the distribution of rights, status, and goods across many social contexts. As data scientists develop tools that inform important social and political decisions—who receives a job offer, what news people see, where police patrols—they shape social outcomes around the world. Data scientists are some of today’s most powerful (and obscured) political actors, structuring how institutions conceive of problems and make decisions.

Why Politics?

This section responds to three arguments that are commonly invoked by data scientists when they are challenged to take political stances regarding their work. These arguments have been expressed in a variety of public and private settings and will be familiar to anyone who has engaged in discussions about the social responsibilities of data scientists.

These are by no means the only arguments proffered in this larger debate, nor do they represent any sort of unified position among data scientists. Nonetheless, the three positions considered here are among the most common and

compelling arguments made against a politically oriented data science. Any promotion of a more politically engaged data science must contend with them.

Argument 1: "I am Just an Engineer"

This first argument represents a common attitude among engineers. In this view, although engineers develop new tools, their work does not determine how a tool will be used. Artifacts are seen as neutral objects that lack any inherent normative character and that can simply be used in good or bad ways. By this logic, engineers bear no responsibility for the impacts of their creations.

It is common for data scientists to argue that the impacts of technology are unknowable. However, by articulating their limited role as neutral researchers, data scientists provide themselves with an excuse to abdicate responsibility for the social and political impacts of their work. When a paper that used neural networks to classify crimes as gang-related was challenged for its potentially harmful effects on minority communities, a senior author on the paper deflected responsibility by arguing, "It's basic research" (Hutson, 2018).

Although it is common for engineers to see themselves as separate from politics, many scholars have thoroughly articulated how technology embeds politics and shapes social outcomes. As political theorist Langdon Winner (1986) describes:

"technological innovations are similar to legislative acts or political foundings that establish a framework for public order that will endure over many generations. For that reason, the same careful attention one would give to the rules, roles, and relationships of politics must also be given to such things as the building of highways, the creation of television networks, and the tailoring of seemingly insignificant features on new machines. The issues that divide or unite people in society are settled not only in the institutions and practices of politics proper, but also, and less obviously, in tangible arrangements of steel and concrete, wires and semiconductors, and nuts and bolts."

Even though technology does not conform to conventional notions of politics, it often shapes society in much the same way as laws, elections, and judicial opinions. In this sense, "the scientific workplace functions as a key site for the production of social and political order" (Jasanoff, 2003). Thus, as with many other types of scientists, data scientists possess "a source of fresh power that escapes the routine and easy definition of a stated political power" (Latour, 1983).

There are many examples of engineers developing and deploying technologies that, by structuring behavior and shifting power, shape aspects of society. As one example, when automobiles were introduced onto city streets in the 1920s, they created chaos and conflict in the existing social order (Norton, 2011). Many cities turned to traffic engineers as "disinterested experts" whose scientific methods could provide a neutral and optimal solution. But the engineers' solution contained unexamined assumptions and values, namely, that "traffic efficiency worked for the benefit of all". As traffic engineers changed the timings of traffic signals to enable cars to flow freely, their so-called solution "helped to redefine streets as motor thoroughfares where pedestrians did not belong". These actions by traffic engineers helped shape the next several decades of automobile-focused urban development in US cities.

Although these particular outcomes could be chalked up to unthoughtful design, any decisions that the traffic engineers made would have had some such impact: determining how to time streetlights requires judgments about what outcomes and whose interests to prioritize. Whatever they and the public may have believed, traffic engineers were never "just" engineers optimizing society "for the benefit of all". Instead, they were engaged in the process—via formulas and signal timings—of defining which street uses should be supported and which should be constrained. The traffic engineers may not have decreed by law that streets were for cars, but their technological intervention assured this outcome by other means.

Data scientists today risk repeating this pattern of designing tools with inherently political characters yet largely overlooking their own agency and responsibility. By imagining an artificially limited role for themselves, engineers create

an environment of scientific development that requires few moral or political responsibilities. But this conception of engineering has always been a mirage. Developing any technology contributes to the particular “social contract implied by building that technological system in a particular form” (Winner, 1986).

Of course, we must also resist placing too much responsibility on data scientists. The point is not that, if only they recognized their social impacts, engineers could themselves solve social issues. Technology is at best just one tool among many for addressing complex social problems (Green, 2019). Nor should we uncritically accept the social influence that data scientists have. Having unelected and unaccountable technical experts make core decisions about governance away from the public eye imperils essential notions of how a democratic society ought to function. As Science, Technology, and Society (STS) scholar Sheila Jasanoff (2006) argues, “The very meaning of democracy increasingly hinges on negotiating the limits of the expert’s power in relation to that of the publics served by technology.”

Nonetheless, the design and implementation of technology does rely, at some level, on trained practitioners. This raises several questions that animate the rest of this article. What responsibilities should data scientists bear? How must data scientists reconceptualize their scientific and societal roles in light of these responsibilities?

Argument 2: “We Shouldn’t Take Political Stances”

Data scientists adhering to this second argument likely accept the response to Argument 1 but feel stuck, unsure how to appropriately act as more than “just” an engineer. “Sure, I am developing tools that impact people’s lives”, they may acknowledge, before asking, “But is not the best thing to just be as neutral as possible?”

Although it is understandable how data scientists come to this position, their desire for neutrality suffers from two important failings. First, neutrality is an unachievable goal, as it is impossible to engage in science or politics without being influenced by one’s background, values, and interests. Second, striving to be neutral is not itself a politically neutral position. Instead, it is a fundamentally conservative one—in the sense that it is committed to maintaining the status quo.

An ethos of objectivity has long been prevalent among scientists. Since the nineteenth century, objectivity has evolved into a set of widespread ethical and normative scientific practices. Conducting good science—and being a good scientist—meant suppressing one’s own perspective so that it would not contaminate the interpretations of observations (Daston & Galison, 2007).

Yet this conception of science was always rife with contradictions and oversights. Knowledge is shaped and bounded by the social contexts that generated it. This insight forms the backbone of standpoint theory, which articulates that “nothing in science can be protected from cultural influence—not its methods, its research technologies, its conceptions of nature’s fundamental ordering principles, its other concepts, metaphors, models, narrative structures, or even formal languages” (Harding, 1998). Although scientific standards of objectivity account for certain kinds of individual subjectivity, they are too narrowly construed: “methods for maximizing objectivism have no way of detecting values, interests, discursive resources, and ways of organizing the production of knowledge that first constitute scientific problems, and then select central concepts, hypotheses to be tested, and research designs” (Harding, 1998).

Every aspect of science is imbued with the characteristics and interests of those who produce it. This does not invalidate every scientific finding as arbitrary, but points to science’s contingency and reliance on its practitioners: all research and engineering are developed within particular institutions and cultures and with particular problems and purposes in mind.

Just as it is impossible to conduct science in any truly neutral way, there is no such thing as a neutral (or apolitical) approach to politics. As philosopher Roberto Unger (1987) writes, political neutrality is an “illusory and ultimately idolatrous goal” because “no set of practices and institutions can be neutral among conceptions of the good.”

Instead of being neutral and apolitical, attempts to be neutral and apolitical embody an implicitly conservative (that is, status quo-preserving) politics. Because neutrality does not mean value-free—it means acquiescence to dominant social and political values, freezing the status quo in place. Neutrality may appear to be apolitical, but that is only because the status quo is taken as a neutral default. Anything that challenges the status quo—which efforts to promote social justice must do by definition—will therefore be seen as political. But efforts for reform are no more political than efforts to resist reform or even the choice simply to not act, both of which preserve existing systems.

Although surely not the intent of every scientist and engineer who strives for neutrality, broad cultural conceptions of science as neutral entrench the perspectives of dominant social groups, who are the only ones entitled to legitimate claims of neutrality. For example, many scholars have noted that neutrality is defined by a masculine perspective, making it impossible for women to be seen as objective or for neutral positions to consider female standpoints (Harding, 1998; Lloyd, 1993; Keller, 1985; MacKinnon, 1982). The voices of Black women are particularly subjugated as partisan and anecdotal (Collins, 2000). Because of these perceptions, when people from marginalized groups critique scientific findings, they are cast off as irrational, political, and representing a particular perspective (Haraway, 1988). In contrast, the practices of science and the perspectives of the dominant groups that uphold it are rarely considered to suffer from the same maladies.

Data science exists on this political landscape. Whether articulated by their developers or not, machine learning systems already embed political stances. Overlooking this reality merely allows these political judgments to pass without scrutiny, in turn granting data science systems with more credence and legitimacy than they deserve.

Predictive policing algorithms offer a particularly pointed example of how striving to remain neutral entrenches and legitimize existing political conditions. The issue is not simply that the training data behind predictive policing algorithms are biased due to a history of overenforcement in minority neighborhoods. In addition, our very definitions of crime and how to address it are the product of racist and classist historical processes. Dating back to the eras of slavery and reconstruction, cultural associations of Black men with criminality have justified extensive police forces with broad powers (Butler, 2017). The War on Drugs, often identified as a significant cause of mass incarceration, emerged out of an explicit agenda by the Nixon administration to target people of color (Alexander, 2012; Baum, 2016). Meanwhile, crimes like wage theft (when employers deny employees wages or benefits to which they are legally entitled) are systemically underenforced by police and do not even register as relevant to conversations about predictive policing—even though wage theft steals more value than other kinds of theft combined (Meixell & Eisenbrey, 2014).

Moreover, predictive policing rests on a model of policing that is itself unjust. Predictive policing software could exist only in a society that deploys vast punitive resources to prevent social disorder, following “broken windows” tactics. Policing has always been far from neutral: “the basic nature of the law and the police, since its earliest origins, is to be a tool for managing inequality and maintaining the status quo” (Vitale, 2017). The issues with policing are not flaws of training or methods or “bad apple” officers, but are endemic to policing itself (Butler, 2017; Vitale, 2017).

Against this backdrop, choosing to develop predictive policing algorithms is not neutral. Accepting common definitions of crime and how to address it may seem to allow data scientists to remove themselves from politics, but instead upholds historical politics of social hierarchy.

Although predictive policing represents a notably salient example of how data science cannot be neutral, the same could be said of all applied data science. Biased data are certainly one piece of the story, but so are existing social and political conditions, definitions and classifications of social problems, and the set of institutions that respond to those problems. None of these factors are neutral and removed from politics. And while data scientists are of course not responsible for creating these aspects of society, they are responsible for choosing how to interact with them. Neutrality in the face of injustice only reinforces that injustice. When engaging with aspects of the world steeped in history and politics, in other words, it is impossible for data scientists to not take political stances.

I do not mean to suggest that every data scientist should share a singular political vision—that would be wildly unrealistic. It is precisely because the field (and world) hosts a diversity of normative perspectives that we must surface political debates and recognize the role they play in shaping data science practice. Nor is my argument meant to suggest that articulating one's political commitments is a simple task. Normative ideals can be complex and conflicting, and one's own principles can evolve over time. Data scientists need not have precise answers about every political question. However, they must act in light of articulated principles and grapple with the uncertainty that surrounds these ideals.

Argument 3: “Don't Let the Perfect Be the Enemy of the Good”

Following the responses to Arguments 1 and 2, data scientists asserting this third argument likely acknowledge that their creations will unavoidably have social impacts and that neutrality is not possible. Yet still holding out against a thorough political engagement, they fall back on a seemingly pragmatic position: because data science tools can improve society in incremental but important ways, we should support their development rather than argue about what a perfect solution might be.

Despite being the most sophisticated of the three arguments, this position suffers from several underdeveloped principles. First, data science lacks robust theories regarding what “perfect” and “good” actually entail. As a result, the field typically adopts a superficial approach to reform that involves making vague (almost tautological) claims about what social conditions are desirable. Second, this argument fails to articulate how to evaluate or navigate the relationship between the perfect and the good. Efforts to promote social good thus tend to take for granted that technology-centric incremental reform is an appropriate strategy for social progress. Yet, considered from a perspective of substantive equality and anti-oppression, many data science efforts to do good are not, in fact, consistently doing good.

Data Science Lacks a Thorough Definition of “Good”

Across the broad world of data science, from academic institutes to conferences to companies to volunteer organizations, “social good” (or just “good”) has become a popular term. While this is both commendable and exciting, the field has not developed (nor even much debated) any working definitions of the term “social good” to guide its efforts. Instead, the field seems to operate on a “know it when you see it” approach, relying on rough proxies such as crime = bad, poverty = bad, and so on. There is, of course, extensive literature (spanning philosophy, STS, and other fields) that considers what is socially desirable, yet data science efforts to promote “social good” rarely reference this literature.

This lack of definition leads to “data science for social good” projects that span a wide range of conflicting political orientations. For example, some work under the “social good” umbrella is explicitly developed to enhance police accountability and promote non-punitive alternatives to incarceration (Bauman et al., 2018; Carton et al., 2016). In contrast, other work under the “social good” label aims to enhance police operations. One such paper aimed to classify gang crimes in Los Angeles (Hutson, 2018; Seo et al., 2018). This project involved taking for granted the legitimacy of the Los Angeles Police Department's gang data—a notoriously biased type of data (Felton, 2018) from a police department that has a long history of abusing minorities in the name of gang suppression (Vitale, 2017). That such politically disparate and conflicting work could be similarly characterized as “social good” should prompt a reconsideration of the core terms and principles. When the term encompasses everything, it means nothing.

The point is not that there exists a single optimal definition of “social good”, nor that every data scientist should agree on one set of principles. Instead, there is a multiplicity of perspectives that must be openly acknowledged to surface debates about what “good” actually entails. Currently, however, the field lacks the language and perspective to sufficiently evaluate and debate differing visions of what is “good”. By framing their notions of “good” in such vague and undefined terms, data scientists get to have their cake and eat it too: They can receive praise and publications based on broad claims about solving social challenges while avoiding substantive engagement with social and political impacts.

Most dangerously, data science's vague framing of social good allows those already in power to present their normative judgments about what is "good" as neutral facts that are difficult to challenge. As discussed earlier, neutrality is an impossible goal and attempts to be neutral tend to reinforce the status quo. If the field does not openly debate definitions of "perfect" and "good", the assumptions and values of dominant groups will tend to win out. Projects that purport to enhance social good but fail to reflexively engage with the political context are likely to reproduce the exact forms of social oppression that many working towards "social good" seek to dismantle.

Pursuing an Incremental "Good" Can Reinforce Oppression

Even if data scientists acknowledge that "social good" is often poorly defined, they may still adhere to the argument that "we should not let the perfect be the enemy of the good". "After all", they might say, "is not some solution, however imperfect, better than nothing?" As one paper asserts, "we should not delay solutions over concerns of optimal" outcomes (Sylvester & Raff, 2018).

At this point the second failure of Argument 3 becomes clear: It tells us nothing about the relationship between the perfect and the good. Although data scientists generally acknowledge that data science cannot provide perfect solutions to social problems, the field typically takes for granted that incremental reforms using data science contribute to the "social good". On this logic, we should applaud any attempts to alleviate issues such as crime, poverty, and discrimination. Meanwhile, because "the perfect" represents an unrealizable utopia we should not waste time and energy debating the ideal solution.

Although efforts to promote "social good" using data science can be productive, pursuing such applications without a rigorous theory of social change can lead to harmful consequences. A reform that seems desirable from a narrow perspective focused on immediate improvements can be undesirable from a broader perspective focused on long-term, structural reforms. Understood in these terms, the dichotomy between the idealized "perfect" and the incremental "good" is a false one: articulating visions of an ideal society is an essential step for developing and evaluating incremental reforms. In order to rigorously conceive of and compare potential incremental reforms, we must first debate and refine our conceptions of the society we want to create; following those ideals, we can then evaluate whether potential incremental reforms push society in the desired direction. Because there is a multiplicity of imagined "perfections", which in turn suggest an even larger multiplicity of incremental "goods", reforms must be evaluated based on what type of society they promote in both the short and long term. In other words, rather than treating any incremental reform as desirable, data scientists must recognize that different incremental reforms can push society down drastically different paths.

When attempting to achieve reform, an essential task is to evaluate the relationship between incremental changes and long-term agendas for a more just society. As social philosopher André Gorz (1961) proposes, we must distinguish between "reformist reforms" and "non-reformist reforms." Gorz explains, "A reformist reform is one which subordinates its objectives to the criteria of rationality and practicability of a given system and policy." In contrast, a non-reformist reform "is conceived not in terms of what is possible within the framework of a given system and administration, but in view of what should be made possible in terms of human needs and demands."

Reformist and non-reformist reforms are both categories of incremental reform, but they are conceived through distinct processes. Reformist reformers start within existing systems, looking for ways to improve them. In contrast, non-reformist reformers start beyond existing systems, looking for ways to achieve emancipatory social conditions. Because of the distinct ways that these two types of reforms are conceived, the pursuit of one versus the other can lead to widely divergent social and political outcomes.

The solutions proposed by data scientists are almost entirely reformist reforms. The standard logic of data science—grounded in accuracy and efficiency—tends toward accepting and working within the parameters of existing systems. Data science interventions are therefore typically proposed to improve the performance of a system rather than to substantively alter it. And while these types of reforms have value under certain conditions, such an ethos of reformist reforms is unequipped to identify and pursue the larger changes that are necessary across many institutions. This approach may even serve to entrench and legitimize the status quo. From the standpoint of existing systems, it is

impossible to imagine alternative ways of structuring society—when reform is conceived in this way, “only the most narrow parameters of change are possible and allowable” (Lorde, 1984).

In this sense, data science’s dominant strategy of pursuing a reformist, incremental good resembles a algorithm pursuing the strategy of making immediate improvements in the **local** vicinity of the status quo. Although this strategy can be useful for simple problems, it is unreliable in complex search spaces: We may quickly find a local maximum but will never reach a further-afield terrain of far better solutions. Moves that are immediately beneficial can be counterproductive for finding the global optimum. Similarly, although reformist reforms can lead to certain improvements, a strategy limited to reformist reforms cannot guide robust responses to complex political problems. Reforms that appear desirable within the narrow scope of a reformist strategy can be counterproductive for achieving structural reforms. Even though the optimal political solution is rarely achievable (and is often subject to significant debate), it is necessary to fully characterize the space of possible reforms and to evaluate how reliably different approaches can generate more egalitarian outcomes.

The US criminal justice system, a domain where data scientists are increasingly striving to do good, exemplifies the limits of a reformist mindset. Because criminal justice reform can be “superficial and deceptive” (Karakatsanis, 2019), it is necessary to couch reform efforts within a broader vision of long-term, non-reformist change. This is the approach taken by the movement for police and prison abolition. Notably, prison abolitionists object to reforms that “render criminal law administration more humane, but fail to substitute alternative institutions or approaches to realize social order maintenance goals” (McLeod, 2013). Instead, abolitionists pursue only reforms that reduce or replace carceral responses to social disorder.

In contrast with this abolitionist ethos, most data science efforts to contribute “good” are grounded in the existing practices of the criminal justice system. A notable example is pretrial risk assessments. Even if they lead to incremental improvements, these tools legitimize policies that drive racial injustice and mass incarceration (Green, 2020). Meanwhile, an entirely separate incremental reform—an abolitionist and non-reformist (and non-technological) one—is possible: ending cash bail and pretrial detention.

Adopting pretrial risk assessments and abolishing pretrial detention appear to respond to the same problems, suggesting that these two reforms are aligned. However, these reforms derive from conflicting visions of the “perfect”. Reformers supporting risk assessments accept pretrial detention as part of criminal justice system, aiming merely to improve the means by which people are selected for pretrial detention. Meanwhile, reformers aiming to abolishing pretrial detention reject pretrial detention, aiming to abolish the practice altogether. In other words, the debate about risk assessments hinges on political questions about how the criminal justice system should be structured. It is only by articulating our imagined perfects that we can even recognize the underlying tension between these two incremental reforms, let alone properly debate which one to pursue.

The point is not that data science is incapable of improving society. However, data science interventions must be evaluated against alternative reforms as just one of many options, rather than compared merely against the status quo as the only possible reform. There should not a default presumption that machine learning provides an appropriate reform for every problem.

In sum, attempts by data scientists to avoid politics overlook technology’s social impacts, privilege the status quo, and narrow the range of possible reforms. The field of data science will be unable to meaningfully advance social justice without accepting itself as political.

References

- Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.
- Anwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. *ProPublica*.
<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

- Baum, D. (2016). Legalize it all. *Harper's Magazine*. <https://harpers.org/archive/2016/04/legalize-it-all/>
- Bauman, M. J., Boxer, K. S., Lin, T.-Y., Salmon, E., Naveed, H., Haynes, L., Walsh, J., Helsby, J., Yoder, S., Sullivan, R., et al. (2018). Reducing incarceration through prioritized interventions. *Proceedings of the 1st ACM SIGCAS conference on computing and sustainable societies*. Association for Computing Machinery.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 81, 77-91.
- Butler, P. (2017). *Chokehold: Policing Black men*. The New Press.
- Carton, S., Helsby, J., Joseph, K., Mahmud, A., Park, Y., Walsh, J., Cody, C., Patterson, E., Haynes, L., & Ghani, R. (2016). Identifying police officers at risk of adverse events. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. Association for Computing Machinery.
- Collins, P. H. (2000). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. Routledge.
- Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books.
- de Montjoye, Y.-A., Radaelli, L., Singh, V. K., & Pentland, A. S. (2015). Unique in the shopping mall: On the reidentifiability of credit card metadata. *Science*, 347(6221), 536-539.
- Felton, E. (2018, March 15). Gang databases are a life sentence for Black and Latino communities. *Pacific Standard*. <https://edtechbooks.org/-WAZQ>
- Gorz, A. (1967). *Strategy for labor*. Beacon Press.
- Green, B. (2021). The contestation of tech ethics: A sociotechnical approach to technology ethics in practice. *Journal of Social Computing*, 2(3), 209-225 doi:10.23919/JSC.2021.0018.
- Green, B. (2020). The false promise of risk assessments: Epistemic reform and the limits of fairness. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 594-606).
- Green, B. (2019). *The smart enough city: Putting technology in its place to reclaim our urban future*. MIT Press.
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
- Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599.
- Harding, S. (1998). *Is science multicultural? Postcolonialisms, feminisms, and epistemologies*. Indiana University Press.
- Hutson, M. (2018, February 28). Artificial intelligence could identify gang crimes—and ignite an ethical firestorm. *Science*. <https://edtechbooks.org/-uYJu>
- Jasanoff, S. (2003). In a constitutional moment: Science and social order at the millennium. In B. Joerges & H. Nowotny (Eds.), *Social studies of science and technology: Looking back, ahead* (p. 155-180). Springer.
- Jasanoff, S. (2006). Technology as a site and object of politics. In R. E. Goodin & C. Tilly (Eds.), *The Oxford handbook of contextual political analysis* (pp. 745-763). Oxford University Press.
- Karakatsanis, A. (2019). The punishment bureaucracy: How to think about "criminal justice reform." *The Yale Law Journal Forum*, 128, 848-935.
- Keller, E. F. (1985). *Reflections on gender and science*. Yale University Press.

- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15), 5802-5805.
- Kramer, A. D. I., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788-8790.
- Lasswell, H. D. (1936). *Politics: Who gets what, when, how*. Whittlesey House.
- Latour, B. (1983). Give me a laboratory and I will raise the world. In K. Knorr-Cetina & M. J. Mulkay (Eds.), *Science Observed: Perspectives on the social study of science* (pp 141-170). Sage.
- Leftwich, A. (1984). Politics: people, resources, and power. In A. Leftwich (Ed.), *What is politics? The activity and its study* (pp. 62-84). Basil Blackwell.
- Lloyd, G. (1993). Maleness, metaphor, and the "crisis" of reason. In L. M. Antony & C. E. Witt (Eds.), *A mind of one's own: Feminist essays on reason and objectivity*. Westview Press.
- Lorde, A. (1984). *Sister outsider: Essays & speeches*. Crossing Press.
- MacKinnon, C. A. (1982). Feminism, Marxism, method, and the state: An agenda for theory. *Signs: Journal of Women in Culture and Society*, 7(3), 515-544.
- McLeod, A. M. (2013). Confronting criminal law's violence: The possibilities of unfinished alternatives. *Unbound: Harvard Journal of the Legal Left*, 8, 109-132.
- Meixell, B., & Eisenbrey, R. (2014, September 11). An epidemic of wage theft is costing workers hundreds of millions of dollars a year. *Economic Policy Institute*. <https://www.epi.org/publication/epidemic-wage-theft-costing-workers-hundreds/>
- Nicas, J. (2018, February 7). How YouTube drives people to the internet's darkest corners. *Wall Street Journal*. <https://edtechbooks.org/-vRPI>
- Norton, P. D. (2011). *Fighting traffic: The dawn of the motor age in the American city*. MIT Press.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
- O'Neil, C. (2017) *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
- Powles, J. (2018, December 7). The seductive diversion of 'solving' bias in artificial intelligence. *Medium*. <https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53>
- Rosenberg, M., Confessore, N., & Cadwalladr, C. (2018, March 17). How Trump consultants exploited the Facebook data of millions. *New York Times*. <https://edtechbooks.org/-RJPW>
- Seo, S., Chan, H., Brantingham, P. J., Leap, J., Vayanos, P., Tambe, M., & Liu, Y. (2018). Partially generative neural networks for gang crime classification with partial information. *Proceedings of the 2018 AAAI/ACM conference on AI, ethics and society (AIES)*. Association for the Advancement of Artificial Intelligence.
- Sylvester & Raff (2018). *What about applied fairness?* Paper presented at The 35th International Conference on Machine Learning. Stockholm, Sweden.
- Thompson, S. A., & Warzel, C. (2019, December 20). How to track President Trump. *New York Times*. <https://edtechbooks.org/-Efhs>

- Unger, R. M. (1987). *False necessity: Anti-necessitarian social theory in the service of radical democracy*. Cambridge University Press.
- Vitale, A. S. (2017). *The end of policing*. Verso Books.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.
- Wexler, R. (2018). Life, liberty, and trade secrets: Intellectual property in the criminal justice system. *Stanford Law Review*, 70(5), 1343-1429.
- Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*. University of Chicago Press.



This content is provided to you freely by EdTech Books.

Access it online or download it at

https://edtechbooks.org/intro_to_data_science/chapter_3_philosophical.

