

Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought

Ben Green
bgreen@g.harvard.edu
Harvard University

Salomé Viljoen
sv486@cornell.edu
Cornell University & New York University

ABSTRACT

Although computer scientists are eager to help address social problems, the field faces a growing awareness that many well-intentioned applications of algorithms in social contexts have led to significant harm. We argue that addressing this gap between the field’s desire to do good and the harmful impacts of many of its interventions requires looking to the epistemic and methodological underpinnings of algorithms. We diagnose the dominant mode of algorithmic reasoning as “algorithmic formalism” and describe how formalist orientations lead to harmful algorithmic interventions. Addressing these harms requires pursuing a new mode of algorithmic thinking that is attentive to the internal limits of algorithms and to the social concerns that fall beyond the bounds of algorithmic formalism. To understand what a methodological evolution beyond formalism looks like and what it may achieve, we turn to the twentieth century evolution in American legal thought from legal formalism to legal realism. Drawing on the lessons of legal realism, we propose a new mode of algorithmic thinking—“algorithmic realism”—that provides tools for computer scientists to account for the realities of social life and of algorithmic impacts. These realist approaches, although not foolproof, will better equip computer scientists to reduce algorithmic harms and to reason well about doing good.

CCS CONCEPTS

- **Theory of computation** → **Design and analysis of algorithms**;
- **Applied computing** → *Law, social and behavioral sciences*.

KEYWORDS

algorithms, law, STS, critical algorithm studies, epistemology

ACM Reference Format:

Ben Green and Salomé Viljoen. 2020. Algorithmic Realism: Expanding the Boundaries of Algorithmic Thought. In *Conference on Fairness, Accountability, and Transparency (FAT* ’20)*, January 27–30, 2020, Barcelona, Spain. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3351095.3372840>

1 INTRODUCTION

Much of the work in computer science labs and technology companies is motivated by a desire to improve society. Many computer scientists aim to “change the world” [113] and contribute to “social

good” [50], leading to the development of algorithms for use in courts [5], city governments [53], hospitals [139], schools [154], and other essential societal institutions. Yet alongside computer science’s growing interest in addressing social challenges has come a recognition—driven by affected communities and scholarship in science, technology, and society (STS) and critical algorithm studies—that many well-intentioned applications of algorithms have led to harm. Algorithms can be biased [5], discriminatory [7], dehumanizing [107], and violent [66, 111]. They can exclude people from receiving social services [42, 109], spread hateful ideas [124, 144], and facilitate government oppression of minorities [89, 103].

Computer science thus faces a gap between its desire to do good and the harmful impacts of many of its interventions. The challenge for the field is how to account for social and political concerns in order to more reliably achieve its aims. Our goal in this paper is to engage with this challenge and to elaborate a positive vision of how computer science can better contribute to society.

To engage in this task, we consider the relationship between “algorithmic thinking” (how algorithms are canonically taught and understood) and “algorithmic interventions” (how algorithms are deployed to address social problems). We are specifically interested in interrogating the influence of “algorithmic thinking” on “algorithmic interventions,” and focus on the application of optimization and machine learning algorithms. The divergent meanings of “algorithms” within critical discourse and computer science [37] reflects the differences between algorithms in theory and the “algorithmic systems—intricate, dynamic arrangements of people and code”—that exist in practice [134]. Understanding the relationship between these two notions of algorithms thus requires approaching algorithms as “multiples”—unstable objects that are enacted through the varied practices that people use to engage with them” [133].

Algorithmic thinking can be understood as a mode of reasoning: it shapes how computer scientists see the world, understand problems, and develop solutions to those problems. Like all methodologies, algorithmics relies on defining the bounds of its analysis. Considerations that fall within a method’s analytic boundaries are the subject of “sharp focus” [131]—but when aspects of the world fall outside these boundaries, a method “has no hope of discovering these truths, since it has no means of representing them” [4]. Thus, as computer science increasingly engages with social and political contexts, the field has come up against the limits of algorithmic thinking: computer science lacks the language and methods to fully recognize, reason about, and evaluate the social aspects and impacts of algorithmic interventions. In turn, even well-intentioned algorithmic interventions are at risk of producing social harms.

Enabling computer science to responsibly navigate its social effects requires several steps: 1) diagnosing the attributes of algorithmic thinking and how those attributes lead to harm, 2) evaluating

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

FAT* ’20, January 27–30, 2020, Barcelona, Spain

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6936-7/20/02...\$15.00

<https://doi.org/10.1145/3351095.3372840>

the potential and limits of current efforts to reform algorithms, 3) describing how the field can expand its epistemic and methodological boundaries, and 4) articulating the tenets of a computer science practice that is evolved based on the concerns raised by affected communities and disciplines such as STS. This paper takes on each of these tasks in turn.

First, we argue that many of the harms of algorithmic interventions derive from the dominant mode of thinking within computer science, which we characterize as “algorithmic formalism.” Algorithmic formalism involves three key orientations: objectivity/neutrality, internalism, and universalism. Although often reasonable (even valuable) within the context of traditional algorithmic work, these orientations can lead to algorithmic interventions that entrench existing social conditions, narrow the range of possible reforms, and impose algorithmic logics at the expense of others. Characterizing these concerns—which draw heavily on STS and critical algorithm studies—under the banner of formalism provides a path to evaluating and pursuing potential remedies.

Second, we evaluate the dominant approaches to reducing algorithmic harms, such as efforts to promote algorithmic fairness, ethics, and various forms of data and model documentation. Such efforts provide important mechanisms to mitigate certain algorithmic harms. Yet these reforms involve incorporating new processes or metrics into the formal method, and thus do not allow practitioners to transcend formalism itself. Additions of form—most notably, algorithmic fairness—fail to provide the epistemic and methodological tools necessary to fully identify and act upon the social implications of algorithmic work. To solve the chronic failures of algorithmic formalism, computer scientists need new modes of reasoning about the social, both as a terrain of intervention and as an attribute of their own work. This requires an evolution of algorithmic reasoning, expanding the bounds of what it means to “think” algorithmically and “do” algorithmic interventions.

Third, we consider a possible path forward. An epistemic and methodological evolution is a daunting task, and it is not obvious how such a shift could occur or that it would be productive. With this in mind, we draw on our characterization of algorithmic formalism to explore a parallel to formalism in another field—law—and to how legal formalism was addressed with a methodological evolution toward legal realism. From around 1860 through the beginning of the twentieth century, American legal thought was characterized by legal formalism: a project to systematize law around scientific and deductive principles. Because this mode of thought adhered to objective principles but did not consider those principles’ actual impacts, its application upheld highly unequal social conditions. These impacts provoked critiques that led to a methodological evolution toward legal realism. Legal realism did not wholly supplant formalism, but instead provided lawyers and judges with additional tools to account for the realities of social life and of the law’s impacts. This shift—which expanded the terrain on which law could be evaluated and debated—suggests both a path toward reforming computer science and the utility of such a path.

Fourth, drawing on the lessons of legal realism, we propose a new mode of computer science thinking—“algorithmic realism”—that responds to the concerns raised by STS and related disciplines. Compared to algorithmic formalism, algorithmic realism provides three alternative orientations: a reflexive political consciousness, a

porousness that recognizes the complexity and fluidity of the social world, and contextualism. As such, algorithmic realism provides the epistemic and methodological tools to develop algorithmic interventions that question unjust social conditions, expand the range of possible reforms, and account for a wide array of values and goals.

At first glance the law may seem like an unusual place to look for inspiration regarding computer science. With a few exceptions [76, 88], law and computer science are typically seen as in tension, or subject to opposing logics: technology moves “fast” while law is “slow,” technology is about “innovation” while law is about “regulation,” and so on. Yet several parallels suggest why this comparison is apt. Algorithmic interventions operate in a manner akin to legal ones, often taking the place of (or, more precisely, offering a particular technical form of) legal reforms. Like the law, algorithms are commonly invoked as neutral mechanisms of formalized decision making. Yet in practice, both are subject to debates regarding the proper role for discretion, ways to combat discrimination, and determinations of the legitimate bases for decision making. Moreover, the recent surge of enthusiasm for “public interest technology” explicitly follows in the footsteps of (and indeed, takes its name from) a prior movement in legal education [129].

Of course, our goal is not to claim a neat one-to-one correspondence between computer science and law (there certainly are substantial differences), but to point to how the lessons of law can inform computer science. Like computer science, the law involves training in a methodological practice that structures how its practitioners create and evaluate social interventions. Modes of legal thought influence legal interventions in much the same way that modes of algorithmic thought influence algorithmic interventions. Legal scholars have long considered the relationship between the intended and actual impacts of social interventions. Thus, we see the parallel to legal formalism/realism as a way to identify a bridge between the deconstructive critique of algorithmic formalism from STS and a new mode of computer science practice—algorithmic realism—that productively engages with these critiques.

Following the history of law, the distinction between algorithmic formalism and realism does not reflect a rigid dichotomy: the evolution toward realism is an *expansion* of computer science to embrace realist orientations alongside formalist ones, not a wholesale *rejection* of formalism. It is precisely the formalism of algorithmic methods that has enabled many of computer science’s most exciting advances [27, 84, 148]. Algorithmic realism provides complementary approaches that make sociotechnical considerations legible and commonplace within computer science thinking. This expanded epistemic and methodological toolkit can help computer scientists to address existing problems more fully and to see new questions.

Nor does the distinction between algorithmic formalism and realism fully characterize the behaviors of computer scientists. In practice, computer scientists are “diverse and ambivalent characters” [133] who blend formalist and realist methods, engaging in “nuanced, contextualized, and reflexive practices” [104] as they “continuously straddle the competing demands of formal abstraction and empirical contingency” [115]. Some computer science subfields (such as CSCW [14]) have long histories of engaging with sociotechnical practices, while others (such as FAT*) are actively developing such methods. We aim to highlight examples of realist-aligned work to help shift such work from exception to standard practice.

Nonetheless, computer scientists recognize that the insights of STS and critical algorithm studies fall beyond their own interpretive frames [100]. Even within the FAT* community, critical evaluations of the mathematization of fairness suggest the need for further evolution from formalism towards realism [11, 52, 58, 67, 135].

Of course, a turn toward algorithmic realism would not remedy or prevent every algorithmic harm. Computer scientists are just one set of actors within larger sociotechnical systems that include other people, institutions, policies, and pressures. Algorithmic realism may do little directly to remedy the harms of algorithms deployed through discriminatory public policies, by authoritarian regimes, or under exploitative business models. A great deal of algorithmic work is also done by people without formal computer science training. Algorithmic thinking presents a potent site for reform, however. Computer science plays an influential role in society, both directly through the work of developing algorithmic interventions and indirectly as algorithmic thinking shapes how scholars (both inside and outside the field), practitioners, and public officials conceive of social challenges and progress [16, 53, 125, 148]. For instance, various public policies and business practices draw on algorithmic reasoning as a way to gain legitimacy [53, 112, 156].

Following STS scholars such as Jasanoff [75] and Winner [149], we aim to trace a middle path between technological determinism and social determinism, exploring the ways in which algorithmic artifacts have politics. We see algorithmic realism not as distinct from sociotechnical systems, but valuable precisely because it situates algorithmic interventions *within* sociotechnical systems. Computer scientists are not the only or the most important actors within sociotechnical systems (nor should they be). Yet reforming such systems requires that computer scientists recognize their positionality and reason about what roles they do (and should) have. Thus, providing computer scientists with the epistemic capacity to navigate the inherent socio-political dimensions of their work is an essential component of sociotechnical reform.

2 ALGORITHMIC FORMALISM

Formalism implies an adherence to prescribed form and rules. The chosen form (e.g., text or numbers) is analyzed according to particular rules, often with the explicit purpose of “constrict[ing] the choice of [a] decisionmaker” or analyst [127]. In literature, for example, formalism involves “the view that the formal properties of a text—its form, structure, grammar, and so forth—define its boundaries” such that “the text stands on its own as a complete entity” [22]. Similarly, formalism in mathematics involves the idea that “mathematics is not a body of propositions representing an abstract sector of reality but is much more akin to a game” where meaning derives from manipulating symbols according to the rules [145].

Formalism is not itself intrinsically bad. It is a method, one that has many virtues. Conceptually, formalizing a problem can lead to analytical clarity. Practically, formalizing a problem can make complex problems tractable. No system that decides how to distribute water, govern educational resources, or predict weather can do so without in some way formalizing the problem at hand.

Yet formalism also has significant limitations. The danger is not in formal methods *per se*, but in the failure to recognize the limits of formal insights and in the uncritical deployment of formal methods

in complex social contexts where their assumptions may be invalid. Because formal knowledge “requires a narrowing of vision,” writes James Scott, “the formal order encoded in social-engineering designs inevitably leaves out elements that are essential to their actual functioning” [131]. This narrowing—who or what is left on the epistemic cutting room floor and systemically excluded, or made the focus of overly-simplified analytical scrutiny—involves political decisions about what is and is not important. Formal orders “are often sites of political and social struggles” with “brutal consequences” for those being classified [15].

Formalism is at the core of algorithms. As one canonical algorithms textbook describes, an algorithm is a “well-defined computational procedure” for solving “a well-specified computational problem” [27]. The essential attribute of this reasoning is formalism through abstraction: algorithms require the explicit mathematical articulation of inputs, outputs, and goals. This process of employing abstraction to “formulat[e] a problem to admit a computational solution” is deemed the hallmark of “computational thinking” [148]. As another introductory algorithms textbook explains, “At their most effective, [...] algorithmic ideas do not just provide solutions to well-posed problems; they form the language that lets you cleanly express the underlying questions” [84]. Done well, “a clean algorithmic definition can formalize a notion that initially seems too fuzzy and nonintuitive to work with mathematically” [84].

Formalism has been a consistent subject of critique in computing. Brian Smith argued that computer scientists should be attentive to the gulf between the abstractions of models and the complexity of the world [136]. Philip Leith called for computer science to replace computing formalism with a “sociological imagination” [87]. Philip Agre decried “the false precision of formalism” as “an extremely constricting” cognitive style [1]. Pat Langley noted that machine learning research has seen an “increased emphasis on mathematical formalization and a bias against papers that do not include such treatments” [86]. More recently, scholars have highlighted the limits of abstraction and formalism with regard to algorithmic fairness, articulating the need for a sociotechnical frame [135].

2.1 Objective and Neutral

The algorithmic formalist emphasis on objectivity and neutrality occurs on two related levels. First, algorithms are perceived as neutral tools and are often argued for on the grounds that they are capable of making “objective” and “neutral” decisions [53, 73, 78, 118]. Second, computer scientists are seen by themselves and others as neutral actors following the scientific principles of algorithm design from positions of objectivity [50, 71]. Such an ethos has long been prevalent among scientists, for whom objectivity—“the suppression of some aspect of the self, the countering of subjectivity”—has become a widespread set of ethical and normative practices [32].

The orientation of objectivity and neutrality prevents computer scientists from grounding algorithmic interventions in explicit definitions of desirable social outcomes. Social and political¹ concerns are granted little space within algorithmic thinking, leaving data science projects to emerge through the negotiation of technical

¹We invoke politics not in the sense of ideologies, political parties, or elections, but (in a manner akin to Winner [149]) to reference broader debates about “the good”—i.e., the set of processes and dynamics that shape social outcomes and that distribute power among people and groups.

considerations such as data availability and model accuracy, with “explicit normative considerations [rarely] in mind” [114]. Even efforts that are motivated as contributions to “social good” typically lack a clear explanation of what such “good” entails, instead relying on vague and undefined notions [50]. Computer scientists engaged in algorithmic interventions have argued, “I’m just an engineer” [71] and “Our job isn’t to take political stances” [50].

This emphasis on objectivity and neutrality leads to algorithmic interventions that reproduce existing social conditions and policies. For objectivity and neutrality do not mean value-free—they instead mean acquiescence to dominant scientific, social, and political values. Scientific standards of objectivity account for certain kinds of individual subjectivity, but “methods for maximizing objectivism have no way of detecting values, interests, discursive resources, and ways of organizing the production of knowledge” [63]. As such, the supposedly objective scientific “gaze from nowhere” is nothing more than “an illusion” [62]. Neutrality similarly represents an “illusory and ultimately idolatrous goal” that often serves to freeze existing conditions in place [141]. Conceptions of algorithms and computer scientists as objective and neutral launder the perspectives of dominant social groups into perspectivelessness, reinforcing their status as the only ones entitled to legitimate claims of neutrality [25, 30, 62, 63]. Anything that challenges existing social structures is therefore seen as political, yet reform efforts are no more political than efforts to resist reform or even the choice simply to not act, both of which preserve existing structures.

Predictive policing systems offer a particularly pointed example of how striving to remain neutral entrenches and legitimizes existing political conditions. These algorithms exist on the backdrop of a criminal justice system that is increasingly recognized as defined by racial injustice. Definitions of crime are the products of racist and classist histories that associated black men with criminality [3, 8, 20, 138]. Moreover, predictive policing is based on the discredited model of “broken windows” policing that has been found to be ineffective and racially discriminatory [20]. In this context, algorithms that uphold common definitions of crime and how to address it are not (indeed, cannot be) removed from politics—they merely *seem* removed from politics. Computer scientists are not responsible for this context, but they are responsible for choosing how they interact with it. When intervening in social contexts steeped in contested histories and politics, in other words, it is impossible for computer scientists to *not* take political stances.

The point is not to argue for a single “correct” conception of the good to which all computer scientists must adhere. It is precisely because a multiplicity of perspectives exists that judgments regarding scientific practices and normative commitments must be explicitly incorporated into algorithm development and evaluation. Yet the epistemic commitments of neutrality and objectivity exclude such considerations from algorithmic reasoning, allowing these judgments to pass without deliberation or scrutiny.

2.2 Internalist

Another attribute of algorithmic formalism is internalism: only considerations that are legible within the language of algorithms—e.g., efficiency and accuracy—are recognized as important design and evaluation considerations. The analysis of an algorithm primarily

emphasizes its run time (or efficiency), characterizing its behaviors in terms of upper, lower, and tight bounds—all features that can be mathematically defined based on the algorithm’s operations [27, 84]. Machine learning algorithms are additionally centered on a corpus of data from which to derive patterns and are evaluated according to accuracy metrics such as area under the curve (AUC). From predictive policing [101] to healthcare [41] to fake news [152], claims regarding an algorithm’s effectiveness and quality emphasize accuracy along these metrics. This approach of defining and measuring algorithms by their mathematical characteristics provides little internal capacity to reason about the social considerations (such as laws, policies, and social norms) that are intertwined with these algorithms’ performance and impacts.

The internalist emphasis on the mathematical features of algorithms leads to algorithmic interventions based in a technologically determinist theory of social change. Because significant aspects of the social and political world are illegible within algorithmic reasoning, these features are held as fixed constants “outside” of the algorithmic system. In turn, algorithms are proposed as a sole mechanism of social change, with the existing social and political conditions treated as static. For instance, several papers analyzing recidivism prediction tools explicitly describe crime rates in this manner. One describes recidivism prevalence as a “constraint—one that we have no direct control over” [24], while another explains, “the algorithm cannot alter the risk scores themselves” [26]. In an immediate sense it is reasonable to see existing recidivism rates as the backdrop against which a risk assessment makes predictions; yet the recurring practice in computer science of treating these social conditions as a fixed “constraint” exemplifies the internalist assumption that algorithms operate atop a static world.

Internalist reasoning leads to algorithmic interventions that optimize social systems according to existing policies and assumptions, drastically narrowing the range of possible reforms. Algorithmic interventions conceived through the internalist orientation have a tendency “to optimize the status quo rather than challenge it” [23]. The goal becomes to *predict* (static) distributions of social outcomes in order to make more informed decisions rather than to *shift* (fluid) distributions in order to enable better outcomes. In this vein, an argument made for risk assessments is that “[a]lgorithms permit unprecedented clarity” because they “let us precisely *quantify trade-offs* among society’s different goals” (e.g., fairness and low crime rates); algorithms thereby “force us to make more explicit judgments about underlying principles” [83]. Yet this calculus can be seen to provide “clarity” only if the contingency and contestability of social conditions are beyond consideration. Note the tradeoffs that fall beyond the purview of risk assessments and are therefore rendered irrelevant: for instance, the tradeoff between pretrial detention and due process or the tradeoff between implementing risk assessments and abolishing pretrial detention.

Moreover, because this internalist orientation emphasizes an algorithm’s mathematical properties, algorithmic interventions are unable to account for the particular ways that people, institutions, and society will actually interact with algorithms. This is one reason why the deployment of algorithms can generate unintended social outcomes. Algorithmic interventions are thus *indeterminate*: the deployment of an algorithm provides little guarantee that the social impacts expected according to internalist evaluations will

be realized. A paradigmatic case involves traffic optimization algorithms, which are modeled on the assumption that increasing road capacity will reduce traffic [108]. Such algorithms have informed urban planning interventions over the past century, from efforts in the 1920s to manage the influx of automobiles [108] to today's visions for self-driving cars [140]. Yet because they rarely account for the second-order effects of their own introduction, these algorithms drastically overestimate the benefits of increasing roadway capacity: in response to more efficient automobile travel, motorists change their behavior to take advantage of the new road capacity, ultimately leading to more driving and congestion [38, 40].

It is impossible for an algorithm to account for every aspect of society or every way that people might respond to it. Every method needs to set boundaries. Yet the choice of where to set those boundaries shapes what factors are considered or ignored, and in turn shapes the impacts of interventions developed through that method [131]. Internalism enforces a strict frame of analysis, preventing algorithmic interventions from adapting to social considerations that are material to success. Computer scientists therefore need to reason more thoroughly about when certain factors can be ignored and when they must be grappled with.

2.3 Universalism

Algorithmic formalism emphasizes an orientation of universalism: a sense that algorithms can be applied to all situations and problems. Popular algorithms textbooks extol the “ubiquitous” applications of algorithms [27] and the “pervasive” reach of algorithmic ideas [84]. An influential computer scientist hails “computational thinking” as “the new literacy of the 21st century,” excitedly describing how this mode of thinking “has already influenced the research agenda of all science and engineering disciplines” and can readily be applied in daily life [148]. While some have recognized that there are contexts in which it is better not to design technology [10], the common practice among computer scientists is to focus on *how* to design algorithms rather than *whether* algorithms are actually appropriate in any given context. In fact, when students in data science ethics classes have questioned whether algorithms should be used to address social challenges, they are told that the question is out of scope (J. Geffert, personal communication, April 4, 2019) [150].

This universalist orientation leads to interventions developed under an assumption that algorithms can provide a solution in every situation—an attitude that has been described in recent years as “technological solutionism” [102], “tech goggles” [53], and “technochauvinism” [18]. Algorithmic interventions have been proposed as a solution for problems ranging from police discrimination [43, 53] to misinformation [152, 156] to depression detection [21, 33]. Numerous initiatives strive to develop data science and artificial intelligence for “social good” across a wide range of domains, typically taking for granted with scant justification that algorithms are an effective tool for addressing social problems [50].

Algorithmic interventions pursued under universalism impose a narrow algorithmic frame that structures how problems are conceived and limits the range of “solutions” deemed viable. Given that “[t]he way in which [a] problem is conceived decides what specific suggestions are entertained and which are dismissed” [34], applying algorithmic thinking to social problems imposes algorithmic

logics—namely, accuracy and efficiency—onto these domains at the expense of other values. In “smart cities,” for instance, algorithms are being deployed to make many aspects of municipal governance more efficient [53]. Yet efficiency is just one of many values that city governments must promote, and in fact is often in tension with those other values. Inefficient behaviors (such as staff making small talk with residents) can improve a municipality’s ability to provide fair social services and garner public trust [151]. More broadly, an emphasis on efficiency in urban life can erode vital civic actions such as deliberation, dissent, and community building [48].

Algorithms can, of course, model a variety of contexts. Efficiency and accuracy are often important factors. But they are typically not the only nor the most important factors. Algorithmic interventions require reasoning about what values to prioritize and what benefits algorithms can provide. However, the universalist orientation prevents computer scientists from recognizing the limits of algorithms and thoroughly evaluating whether algorithms are appropriate.

This uncritical deployment of algorithmic interventions in turn elevates the status of the algorithmic reasoning behind such interventions. Algorithmic formalism has in many ways become the hallmark of what it means to conceive of any problem rigorously, regardless of the many examples of how such thinking faces serious epistemic defects in various social settings. As such, a significant risk of algorithmic formalism is that it contributes to formal methods dominating and crowding out other forms of knowledge and inquiry (particularly local forms of situated knowledge) that may be better equipped to the tasks at hand.

3 FORMALIST INCORPORATION

One approach to addressing the failures of algorithmic formalism is to incorporate new processes, variables, or metrics into its logic. This process, which we call “formalist incorporation,” is particularly appealing to practitioners operating within algorithmic formalism, who tend to respond to critiques of formalizations with calls for alternative formalizations [1]. For example, one paper that describes an algorithmic intervention whose implementation was blocked by community resistance notes, “the legitimate concerns raised by these families can be modeled as objectives within our general formulation and integrated within our framework” [12].

We see many of the recent efforts in the algorithm research and policy communities as examples of formalist incorporation. Specific interventions of this sort include the methods of algorithmic fairness and approaches to improve data and model documentation [46, 69, 99]. Such reforms have significant value and can improve many aspects of algorithms, but they are not designed to provide an alternative mode of reasoning about algorithms. Similarly, although the burgeoning frame of ethics has potential to expand algorithmic reasoning, efforts to promote ethics within computer science and the tech industry have tended to follow a narrow logic of technological determinism and technological solutionism [44, 59, 96, 150]. Because these reforms operate within the logic of algorithmic formalism, they are ultimately insufficient as remedies: formalist incorporation cannot address the failures of formalism itself. When computer scientists raise concerns and engage with social science in this manner, “broader epistemological, ontological, and political questions about data science tools are often sidelined” [100].

We focus here on the methods and research regarding algorithmic fairness, which represents (arguably) the most significant recent change to algorithmic research and practice in response to algorithmic harms. As currently conceived, algorithmic fairness is ill-equipped to address these concerns because it is itself a manifestation of algorithmic formalism via formalist incorporation.

First, algorithmic fairness is grounded in objectivity and neutrality. Fairness is treated as an objective concept, one that can be articulated and pursued without explicit normative commitments [58]. Approaches to algorithmic fairness often position their goals “in painfully neutral terms” such as “non-discrimination” [67]. Much of the work on fairness points to “bad actors,” [67], reinforcing the view that algorithms themselves are neutral. In turn, emphasizing an algorithm’s fairness often obscures deeper issues such as unjust practices and policies [53]. What may appear “fair” within a narrow computational scope can reinforce historical discrimination [54].

Second, fairness relies on a narrow, internalist approach: “the mandate within the fair-ML community has been to mathematically define aspects of the fundamentally vague notions of fairness in society in order to incorporate fairness ideals into machine learning” [135]. For example, one paper explicitly “reformulate[s] algorithmic fairness as constrained optimization” [26]. The deployment of an algorithm mathematically deemed “fair” is assumed to increase the fairness of the system in which the algorithm is embedded. For example, predictive policing algorithms and risk assessments have been hailed as remedying the injustices of the criminal justice system [43, 53, 117, 123, 143], with significant energy spent ensuring that these algorithms satisfy mathematical fairness standards. Yet such assessments typically overlook the ways in which these “fair” algorithms can lead to unfair social impacts, whether through biased uses by practitioners [2, 29, 55, 56], entrenching unjust policies [54], distorting deliberative processes [51], or shifting control of governance toward unaccountable private actors [17, 77, 146].

Third, fairness embodies an attitude of universalism. Attempts to define and operationalize fairness treat the concept universally, with little attention to the normative meaning behind these definitions or to the social and political context of analysis [58, 135]. Much of the algorithmic fairness literature prioritizes portability of definitions and methods across contexts [135]; evaluation tools are designed to fit into any machine learning pipeline [126]. In turn, fairness is applied as the solution wherever algorithmic biases (or other harms) are exposed. For instance, when research exposed that face and gender recognition systems are more accurate on light-skinned men than on dark-skinned women [19], the primary response was to strive for less biased systems [85, 95, 122, 147, 155], in one case by targeting homeless people of color for facial images [112]. Yet such a pursuit of fair facial recognition does not prevent the systemic harms of this technology—instead, making facial recognition “fair” may legitimize its use under the guise of technical validation [65].

Because of its formalist underpinnings, fair machine learning fails to provide the tools for computer scientists to engage with the critical normative and political considerations at stake when developing and deploying algorithms. Addressing the ways in which algorithms reproduce injustice requires pursuing a new mode of algorithmic thinking that is attentive to the social concerns that fall beyond the bounds of algorithmic formalism.

4 METHODOLOGICAL REFORM: FROM FORMALISM TO REALISM IN THE LAW

To understand the nature and impacts of an intervention to remedy the limits of formalist reasoning, we turn to the evolution in American legal thought from legal formalism to legal realism.

4.1 Legal Formalism

The period from about 1860 through the First World War was one of consensus in American legal thought. The dominant method, called “legal formalism,” was the product of concerted effort by legal scholars and judges to promote formal methods in law [81].² Legal formalism provided a both a descriptive and a normative account, positing how judicial reasoning does and should occur.

American legal thought in this period was “formal” in several senses. Most fundamentally, law was seen “as a science [that] consists of certain principles or doctrines” [130]. Legal thought aimed to identify, classify, and arrange the principles embodied in legal cases as part of a unified system. Jurists working in this mode tended to see legal authority as separated along “sharp analytical boundaries—between public and private, between law, politics and morality, and between state and civil society” [82]. Each entity exercised absolute power within its sphere of authority but was not supposed to consider what lay beyond its internalist bounds. Legal formalists favored the application of law along a series of “bright-line” rules; these rigid rules were believed to create a more objective and scientific application of law because they prevented exceptions or context-specific claims. Finally, legal formalism aspired to determinism. It was assumed that a small number of universal principles, derived from natural rights, could be applied to reliably deduce the correct application of law in specific instances [81, 82].

The height of legal formalism coincided with the period of *laissez-faire* policies and provided reasoning well-suited to defend these policies from progressive challenge. Legal formalism emphasized the autonomy of private citizens and the divide between the authority of the state and that of private actors. From these general principles, judges deduced that efforts to regulate the economy were unconstitutional [142]. In the seminal 1905 case *Lochner v. New York*, the U.S. Supreme Court concluded that a law limiting the working hours of employees represented “unreasonable, unnecessary and arbitrary interference with the right and liberty of the individual to contract” [142]. In his dissent, Justice Oliver Wendell Holmes argued that the Court failed to consider the context of the case, noting, “General propositions do not decide concrete cases” [142]. Legal scholar Roscoe Pound argued that *Lochner* reflected an ignorance of actual working conditions in the United States, which he attributed in part to the “blinders imposed on judges by the ‘mechanical’ style of judicial reasoning” [120]. Following *Lochner*, it became clear among reform-minded legal scholars that enabling the law to account for the realities of social life necessitated, as a first step, methodological critiques of the formal reasoning that judges used to uphold the status quo.

²The term “legal formalism” was not used by its adherents but was first introduced by legal realists to describe the dominant mode of reasoning they sought to displace. Contemporary legal scholars typically refer to this mode of reasoning (and the period in which it was dominant) as Classical Legal Thought.

4.2 Legal Realism

The consensus around legal formalism was upended by an alternative mode of thought: “legal realism.” Motivated by what they saw as the failure of legal reasoning to account for its real-world impacts, the legal realists challenged the formalist “jurisprudence of forms, concepts and rules” [82]. They believed that the inability of supposedly well-reasoned legal analysis to address social challenges such as poor working conditions and staggering inequality stemmed from the fact that context-specific realities and the social impact of laws had no place in formal legal analysis.

Achieving social reform therefore required a methodological intervention: a shift in the everyday reasoning of lawyers and judges in order to render social concerns legible in legal thought. Holmes wrote that the “main purpose” of legal realist interventions “is to emphasize certain oft-neglected matters that may aid in the understanding and in the solution of practical, every-day problems of the law” [68]. This pragmatic approach to reform was deeply rooted in the commitment of the legal realists to create a “realistic jurisprudence” focused not on the “paper rules” of black letter doctrine, but the “real rules” that actually described the behavior of courts [90]. Realists aimed to enable the law (and themselves as practitioners of the law) to deal “with things, with people, with tangibles [...]—not with words alone” [91]. Rather than simply point out the failures of legal formalism, realist critiques put forward new modes of practical reasoning that overcame the epistemic limitations of formalism and that expanded the commonsense modes of “thinking like a lawyer.”

4.2.1 From Universal Principles to Contextual Grounding. A primary legal realist insight was that legal outcomes were not—and could not be—the result of a scientific process. Wesley Hohfeld argued that formal legal thought engaged in deductive error by treating legal principles as universal: “the tendency—and the fallacy—has been to treat the specific problem as if it were far less complex than it really is; and this [...] has [...] furnished a serious obstacle to the clear understanding, the orderly statement, and the correct solution of legal problems” [68].

The issue arose because efforts to deduce rights and duties from universal principles of liberty or autonomy overlooked how the law was indeterminate (e.g., it could protect “liberty” in multiple competing yet equally plausible ways), confronting decision makers with “a choice which could not be solved by deduction” [82]. In conventional examples of legal reasoning, legal realists identified instances of “legal pluralism”—the capacity for legal materials (e.g., prior cases, statutes, rules and principles) to render multiple legitimate outcomes due to gaps, conflicts, ambiguities, and circularities within those materials. The resulting indeterminacy and pluralism forced legal actors to make judgments, based on their interpretations and values rather than mechanical procedures, that would structure social dynamics—in effect, making policy.

Realists argued that this adherence to deduction from general principles played a key role in law’s complicity with the social harms of the day. Formal legal analysis was evaluated based on whether it correctly identified and applied legal principles. This privileged the formally correct application of principles over the (often unequal) results that such applications created. Realists decried

this adherence to “an academic theory of equality in the face of practical conditions of inequality” as methodologically absurd as well as socially harmful [120]. Instead, realists asserted, legal decisions should be evaluated based on their actual impact in their particular context: law should be understood as a means to achieving social ends, not as an end in itself [91]. Unlike philosophy, argued Holmes, the law was not a project of the ideal, but an instrumental means of administering justice in the messy and complex world [70].

4.2.2 From Objective Decisions to Political Assessments. Because cases could not be solved by applying general principles, realists argued that it is impossible to engage in legal decision making without exercising some degree of subjective judgment. The act of filling gaps in legal reasoning with policymaking was thus infused with politics—the ideological predilections and commitments of the judge. The upshot for realists was not that such expressions of politics are inappropriate, but that they are inevitable.

Realist insights enabled legal practitioners to grapple with the policymaking nature of their work. For example, in cases regarding workers’ rights following *Lochner*, judges could no longer reason that judicial interference would be impermissible, because judicial restraint was as much of a political choice as judicial intervention [61]. More broadly, realists displaced the dominance of bright-line rules³ with a shift towards standards meant to structure reasoning regarding law’s social context and impacts.⁴ Moving from rules to standards—rendering gaps in deductive legal reasoning more explicit and legible—was one way that the law evolved its methods to incorporate social context and impact into legal doctrines.

4.2.3 From Internalist Boundaries to Porous Analysis. The effort to evaluate the law vis-à-vis its social impact opened up legal analysis to the languages and methods of other disciplines. Legal realists were enthusiastic about filling normative legal gaps with pragmatist philosophy, political science, statistics, sociology, and economics, and decried the failure of law to keep up with developments in “social, economic and philosophical thinking” [121]. They developed limits to legal reasoning *within* legal authority, carving out spaces where law should defer to these disciplines rather than to a judge.

This emphasis on social impact affected legal analysis in two important ways. First, it opened up terrain for the positive program of incorporating “considerations of social advantage” [70] into legal decision making—to resolve ambiguities in legal materials by looking to social realities. Robert Hale’s analysis of industrial workplace conditions typifies this approach [61]. Hale argued that judicial decisions relying on broad commitments to freedom (and opposition to government coercion) to protect “freedom of contract” from workplace unionization flew in the face of Industrial Era workplace conditions, where workers faced extreme coercive pressure from private employers. Hale showed how legal decisions necessarily distribute freedom and coercion among parties, thus necessitating that decisions be made in reference to a broader social objective.

Second, the focus on impact shifted legal thinking toward considering how opinions and laws would play out in practice. Holmes

³E.g., “If you are on the property of another without consent, they are not liable for any injury you may suffer under trespass.”

⁴E.g., “Under certain conditions, it may be socially desirable for us to enforce liability even under conditions of trespass: for example, if the harm came to a child lured onto the property by an attractive nuisance.”

argued that legal inquiry should concern itself with the messy administration of justice among real-world actors. Legal scholars and judges should therefore think of law as would “a bad man” who is not motivated by “the vaguer sanctions of conscience” but only the “material consequences” that may befall him if he runs afoul of the law [70]. To assess whether a law is good or bad, in other words, legal thinkers ought to anticipate the behavior of actors looking to take advantage of the law.

4.2.4 The Realist Evolution of Legal Common Sense. Realist critiques and proposals were controversial and spurred intense debate [82, 91]. Moreover, realist interventions did not provide a silver bullet to the intractable challenge of administering justice through law. Nor did legal realism fully supplant legal formalism: many formalist orientations remain common in American legal thought (and have in recent decades regained prominence in many areas).

Nonetheless, legal realism provided the methodological basis for profound legal reform. Realist methods enabled progressive changes in private law, provided the intellectual foundations for the administrative state, and led to the overturn of *Lochner v. New York* and the subsequent creation of American labor law [81]. Perhaps legal realism’s most significant contribution was expanding the epistemic and methodological terrain on which legal reasoning and legal debate could occur. By the 1950s, law students became adept at reasoning about the limitations of law and at making arguments about the policy effects of legal decisions. American legal pedagogy “deeply absorbed the basic idea that the validity of laws should be measured, in part, in terms of their social and economic effects” [98]. Realist methods remain highly influential and have provided the intellectual foundation for several subsequent and ongoing efforts to expand legal thinking, including critical legal studies [80], critical race theory [31], law and economics [119], and law and political economy [60].

5 ALGORITHMIC REALISM

Recognizing the dangers of algorithmic formalism and the lessons of legal realism, we turn now to articulating the principles of algorithmic realism. These aspirational attributes counter the orientations of algorithmic formalism, with particular attention to preventing (or at least mitigating) the harms it can produce. As the case of legal thought demonstrates, such a shift can productively enhance a discipline’s epistemic and methodological ability to engage with the social. While no mode of reasoning can avoid imposing its logic on the world, self-conscious modes can expand their internal logic to explicitly reason about their effects on the world.

5.1 Political

Rather than strive for unattainable notions of objectivity and neutrality, algorithmic realism emphasizes that algorithmic interventions are inherently political. This does not entail computer science entirely abandoning objectivity and its practices, such as the norm against manipulating data in order to generate desired results. Instead, it means interrogating the types of subjectivity that typically fly under the radar of “objective” practice: choices such as formulating research questions, selecting methodologies and evaluation metrics, and interpreting results.

This political orientation enables computer scientists to reflect on the normative commitments and outcomes of algorithmic interventions. Rather than creating paralysis, with computer scientists unsure how to be neutral and objective when doing so is impossible, algorithmic realism provides a language to reason about political commitments and impacts as part of what it means to “do” algorithms. First, freed from the strict imperative to be neutral and objective, computer scientists can interrogate the ways in which their assumptions and values influence algorithm design. This reflexive turn can help computer scientists—regardless of their particular normative commitments—better reason about the relationship between their design choices, their professional role, and their vision of the good. Such reflection should occur through open discussion and deliberation, forming a central component of the research process. Second, algorithmic realism shifts the primary focus of algorithmic interventions from the quality of an algorithmic system (in an internalist sense) to the social outcomes that the intervention produces in practice. No matter how technically advanced or impressive a system is, its success under an algorithmic realist frame is defined by whether that system actually leads to the desired social changes.

This approach enables interventions that question rather than uphold unjust social conditions and policies. Several approaches can inform such development of algorithms. The schema of “reformist” and “non-reformist” reforms, articulated by social philosopher André Gorz, provides a way to evaluate interventions based on their political implications [49]. While a reformist reform “subordinates its objectives to the criteria of rationality and practicability of a given system and policy,” 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.” Designers Anthony Dunne and Fiona Raby classify design into two categories: affirmative design, which “reinforces how things are now,” and critical design, which “rejects how things are now as being the only possibility” [39]. A related framework is “anti-oppressive design,” which orients “the choice of a research topic, the focus of a new social enterprise, or the selection of clients and projects” around challenging oppression [137]. Similarly, the Design Justice Network provides ten design principles that include “prioritize design’s impact on the community over the intentions of the designer” [105].

These frameworks show that recognizing algorithmic interventions as political does not prevent computer scientists from doing computer science—instead, doing so can help them incorporate normative reflection into the methods and questions that drive their work. With this in mind, computer scientists can ask a variety of questions to inform their practice: Would the implementation of this algorithm represent a reformist or non-reformist reform? Is the design of this algorithm affirmative or critical? Would providing our project partner with this algorithm entrench or challenge oppression? Is the project prioritizing outcomes over my intentions? Will this algorithm empower the communities it affects?

An example of the expanded practical reasoning that a political orientation provides involves burgeoning activism among employees of technology companies against developing algorithmic interventions for use by the United States Departments of Defense and Homeland Security [47]. Rather than perceiving themselves as “just

an engineer” [71], these computer scientists recognize their position within larger sociotechnical systems, perceive the connection between developing an algorithmic intervention and the political and social outcomes of those interventions, and hold themselves (and their companies) accountable to the impacts of the algorithms they develop. Building on this movement, in 2019, thousands of computer science students from more than a dozen U.S. universities pledged that they would not work for Palantir due to its partnerships with Immigration and Customs Enforcement (ICE) [97].

5.2 Porous

Recognizing algorithmic formalism’s limited ability to characterize sociotechnical systems, algorithmic realism is porous, expanding the range of considerations deemed relevant to algorithm design and evaluation. Factors that were previously beyond the internalist algorithmic frame become central to what it means to have knowledge or make claims about algorithms. A porous approach to algorithms means that formalist considerations (e.g., accuracy, efficiency, and fairness) are recognized as necessary but *no longer sufficient* to define the efficacy or quality of an algorithm—additional modes of analysis are essential. As in law, realism entails both an appreciation of the insights of other fields and a willingness, where appropriate, to carve out spaces of deference to those fields.

This porous orientation allows for algorithmic interventions that eschew technological determinism and instead recognize the contingency and fluidity of the social world. It makes legible the potential for social and policy change in addition to (or instead of) technological change. This does not mean adopting a mantra of social determinism, believing that social systems will evolve irrespective of technology. Instead, a porous approach to algorithmic interventions follows an STS understanding of how “the realities of human experience emerge as the joint achievements of scientific, technical and social enterprise” [75].

This porous orientation gives computer scientists the capacity to widen rather than narrow the range of possible reforms. Rather than optimizing existing systems under the assumption of a static society, computer scientists can develop interventions under the recognition of a fluid society. Several projects exemplify this approach. For example, instead of developing predictive policing or risk assessment algorithms that treat risk levels and policy responses as static, computer scientists have developed algorithms to reduce the risk of crime and violence through targeted and non-punitive social services [9, 57]. In other contexts, computer scientists have subordinated their priorities to broader communities, helping to empower groups advocating for change [6, 28, 35, 79, 93, 94, 128].

Furthermore, by bringing the social world into the algorithmic frame, a porous orientation allows for algorithmic interventions that recognize and account for indeterminacy. Under algorithmic realism, “good” algorithm design means not simply designing to promote desired outcomes, but defining what outcomes are desirable and undesirable, understanding how potential harms could arise, and developing anticipatory mechanisms to prevent or mitigate those outcomes. By incorporating these considerations as essential to algorithm design, algorithmic realism casts practices such as failing to consider how users interact with an algorithm as no less negligent than failing to test a model’s accuracy.

Although it is impossible to fully account for indeterminacy or to guarantee that an intervention will have particular impacts, scholarship from STS and critical algorithm studies provides valuable starting points for analyzing the relationship between algorithmic interventions and social impacts. The Social Construction of Technology (SCOT), for example, argues that new technologies contain numerous potential interpretations and purposes; how a technology stabilizes (in “closure”) depends on the social groups involved in defining that technology and the relative resources each has to promote its particular vision [116]. Co-production more richly articulates the intertwined nature of technology and social conditions, noting identities, institutions, discourses, and representations as particularly salient pathways of social and technological change [75]. A great deal of other recent work has documented the particular ways in which the design, application, and use of algorithms can exacerbate marginalization and inequality [18, 42, 53, 64, 107, 110].

Taking these approaches as a guide, numerous questions can inform computer scientists’ understanding of how an algorithm will interact with and impact communities in a given context. These include: Who are the relevant social actors? What are their interests and relative amounts of power? Which people need to approve this algorithm? What are their goals? On whose use of the algorithmic system does success depend? What are their interests and capabilities? How might this algorithm affect existing scientific, social, and political discourses or introduce new discourses?

This approach has particular value in anticipating and preventing harmful social impacts of algorithms. Just as Holmes urged legal scholars and judges to evaluate laws in light of how they will be carried out in practice, so too should computer scientists evaluate algorithmic interventions through the lens of how people may actually apply them. For example, recognizing how police use of algorithms can distort interventions toward surveillance and punishment, some researchers developing algorithms to identify people at risk of involvement in crime or violence explicitly articulate their commitment to partnering with community groups and social service providers rather than with law enforcement [9, 45, 57].

5.3 Contextual

In contrast to the universalism of algorithmic formalism, algorithmic realism is grounded in contextualism, emphasizing the need to understand social contexts in order to determine the validity of any algorithmic intervention. Rather than question *how* a situation can be modeled and acted upon algorithmically, a contextual approach questions *to what extent* a situation can be modeled and should be acted upon algorithmically. Context is defined here not in a positivist sense of data that can be incorporated into algorithms, but in a broader sense entailing the social relations, activities, and histories that shape any particular setting [36]. Gleaning context therefore requires a porous approach rather than an internalist focus on data [16, 36, 45, 132]. Such context is essential to understanding relationships and behaviors in sociotechnical systems [92, 106].

A contextual orientation allows computer scientists to avoid solutionism and instead take an agnostic approach to algorithmic interventions. Agnosticism entails approaching algorithms instrumentally, recognizing them as just one type of intervention, one that cannot provide the solution to every problem. In other words, an

agnostic approach prioritizes the social impacts of reform, regardless of the role played by algorithms—it is agnostic as to the means, but not the ends. This approach can help not just to avoid harmful algorithms, but also to place algorithms alongside institutional and policy reforms in order to robustly promote well-articulated social ends. For even in contexts where algorithms can help to address social challenges, they cannot do so in isolation: the most impactful algorithmic interventions occur when algorithms are deployed in conjunction with policy and governance reforms [53].

This approach also allows algorithmic thinking to be incorporated into social and policy reform efforts without requiring the deployment of an algorithm and the imposition of algorithmic logics. Contextualism makes legible questions about whether algorithms can capture the essential aspects of a real-world context and whether algorithms can generate the desired social impacts. Computer scientists pursuing interventions through a contextual approach can pose numerous questions: What elements of this context does an algorithmic approach capture and overlook? What values are important for any solution? To what extent can an algorithm account for those values? How does an algorithm compare to other reforms in terms of producing better outcomes? If the answers to these questions suggest a significant divide between the context and an algorithm's ability to model and improve that context, then it is likely that an algorithmic intervention is an ill-advised approach to providing the desired social benefits.

To see this in practice, consider the experience of one of this paper's authors while working as a data scientist with a municipal Emergency Medical Services (EMS) department. The author was asked to improve ambulance response times with data analytics. The instinct of an algorithmic formalist, following a universalist orientation, would be to develop an algorithm that optimizes ambulance dispatch [13, 72, 74, 153]. Yet when the author studied the context of the problem, it became clear that such a "solution" would not fit into EMS's operations nor would it address the underlying issues generating long response times. The author's analysis revealed that significant resources were being deployed to 911 calls for people struggling with homelessness, mental illness, and drug addiction. These individuals did not require the acute medical care that EMS was providing (at the expense of providing it for other incidents); instead they needed social services that EMS was ill-equipped to provide. It became clear that ambulance response efficiency was a limited frame for understanding (and thus reforming) EMS's operations: the efficiency of ambulance responses said nothing about the broader goal of providing services that address people's needs.

Although a dispatch optimization algorithm may perform well along formalist metrics of efficiency, such an algorithm would have failed to address the underlying issue. The author instead worked with EMS to create a new unit of EMTs who would respond to these incidents via bicycle or car and be specially trained to connect people to local social services; the parameters of when and where this unit would operate were determined by analyzing EMS incident data. Notably, the ultimate intervention was *not* to integrate an algorithm into existing procedures: a policy change informed by data was better suited to improve both efficiency and service quality. Rather than representing a failure to take advantage of algorithms, this effort was recognized as a positive collaboration that integrated data analysis and institutional context to improve social services.

6 DISCUSSION

The numerous and significant harms of algorithms may appear to be the result of computer scientists failing to follow best practices. Yet our articulation of algorithmic formalism describes how these outcomes are due to the logic of algorithmic thinking itself, not an imperfect or malevolent application thereof. The chronic tunnel vision of algorithmic formalism can lead to harmful outcomes despite good intentions and following current best practices. Remediating these failings requires not incorporating new variables or metrics (such as fairness) into the formal method but instead introducing new epistemic and methodological tools that expand the bounds of what it means to "do" algorithms.

Algorithmic realism represents this evolution in algorithmic thought, providing new modes of practical reasoning about the relationship between algorithms and the social world. The realist orientations described here provide important starting points for computer scientists and others pursuing algorithmic interventions. Following the political orientation, practitioners should consider what assumptions and values they may be taking for granted and what normative commitments they want their intervention to embody. Following the porous orientation, practitioners should consider what theory of change motivates their work and how to responsibly account for unexpected impacts. Following the contextual orientation, practitioners should consider what goals are central to a given context and whether an algorithm actually provides an appropriate intervention. In a realist mode of reasoning, all of these questions are seen as integral to rigorous algorithmic work rather than as beyond the scope of algorithmic design. These realist practices will enable the field not just to avoid harmful impacts, but also to identify new research questions and directions to pursue.

As in law, algorithmic realism is not meant to provide a wholesale rejection of formal methods nor will it provide a wholesale solution to the intractable challenges of designing just algorithmic systems. Even to the extent that the turn to algorithmic realism is motivated by a broader program of social reform (à la the turn to legal realism), new epistemic and methodological tools cannot by themselves achieve a vision of the good, let alone determine which vision of the good to work towards. Nonetheless, algorithmic realism can help computer scientists reflexively approach their work in light of their larger normative commitments and the impacts of algorithmic systems. As such, algorithmic realism enables computer scientists to reason well about doing good.

ACKNOWLEDGMENTS

We are grateful to the FAT* reviewers and track chair, the Cornell Digital Life Initiative, the NYU Privacy Reading Group, the NYU Information Law Fellows, and participants of PLSC Europe for their thoughtful comments on earlier versions of this paper. We also thank Sheila Jasanoff for prompting the discussions that led to this paper. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE1745303. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

- [1] Philip E. Agre. 1997. Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI. In *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*, Geoffrey C. Bowker, Susan Leigh Star, William Turner, and Les Gasser (Eds.).
- [2] Alex Albright. 2019. If You Give a Judge a Risk Score: Evidence from Kentucky Bail Decisions. *The John M. Olin Center for Law, Economics, and Business Fellows' Discussion Paper Series* 85 (2019).
- [3] Michelle Alexander. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. The New Press.
- [4] Elizabeth Anderson. 2009. Toward a Non-Ideal, Relational Methodology for Political Philosophy: Comments on Schwartzman's "Challenging Liberalism". *Hypatia* 24, 4 (2009), 130–145. www.jstor.org/stable/20618184
- [5] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine Bias. *ProPublica* (2016). <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>
- [6] Mariam Asad. 2019. Prefigurative Design As a Method for Research Justice. *Proceedings of the ACM on Human-Computer Interaction*, 3, CSCW (2019), 200:1–200:18. <https://doi.org/10.1145/3359302>
- [7] Solon Barocas and Andrew D. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104 (2016), 671–732.
- [8] Dan Baum. 2016. Legalize It All. *Harper's Magazine* (2016). <https://harpers.org/archive/2016/04/legalize-it-all/>
- [9] Matthew J. Bauman, Kate S. Boxer, Tzu-Yun Lin, Erika Salomon, Hareem Naveed, Lauren Haynes, Joe Walsh, Jen Helsby, Steve Yoder, and Robert Sullivan. 2018. Reducing Incarceration through Prioritized Interventions. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies (COMPASS '18)*. ACM, 6:1–6:8. <https://doi.org/10.1145/3209811.3209869>
- [10] Eric P.S. Baumer and M. Six Silberman. 2011. When the implication is not to design (technology). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, 2271–2274. <https://doi.org/10.1145/1978942.1979275>
- [11] Sebastian Benthall. 2018. Critical reflections on FAT* 2018: a historical idealist perspective. *DATACTIVE* (2018). <https://data-activism.net/2018/04/critical-reflections-on-fat-2018-a-historical-idealist-perspective/>
- [12] Dimitris Bertsimas, Arthur Delarue, and Sebastien Martin. 2019. Optimizing schools' start time and bus routes. *Proceedings of the National Academy of Sciences* 116, 13 (2019), 5943–5948. <https://doi.org/10.1073/pnas.1811462116>
- [13] Justin J. Boultier and Timothy C.Y. Chan. 2018. Ambulance Emergency Response Optimization in Developing Countries. *arXiv preprint arXiv:1801.05402* (2018).
- [14] Geoffrey Bowker, Susan Leigh Star, Les Gasser, and William Turner. 2014. *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*. Routledge.
- [15] Geoffrey C. Bowker and Susan Leigh Star. 2000. *Sorting Things Out: Classification and Its Consequences*. MIT Press.
- [16] danah boyd and Kate Crawford. 2012. Critical Questions for Big Data. *Information, Communication & Society* 15, 5 (2012), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- [17] Robert Brauneis and Ellen P. Goodman. 2018. Algorithmic Transparency for the Smart City. *The Yale Journal of Law & Technology* 20 (2018), 103–176.
- [18] Meredith Broussard. 2018. *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press.
- [19] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, A. Friedler Sorelle and Wilson Christo (Eds.), Vol. 81. PMLR, 77–91. <http://proceedings.mlr.press>
- [20] Paul Butler. 2017. *Chokehold: Policing Black Men*. The New Press.
- [21] Fidel Cacheda, Diego Fernandez, Francisco J. Novoa, and Victor Carneiro. 2019. Early Detection of Depression: Social Network Analysis and Random Forest Techniques. *Journal of Medical Internet Research* 21, 6 (2019), e12554. <https://doi.org/10.2196/12554>
- [22] Mary Ann Cain. 1999. Problematizing Formalism: A Double-Cross of Genre Boundaries. *College Composition and Communication* 51, 1 (1999), 89–95.
- [23] Nicholas Carr. 2014. The Limits of Social Engineering. *MIT Technology Review* (2014). <https://www.technologyreview.com/s/526561/the-limits-of-social-engineering/>
- [24] Alexandra Chouldechova. 2017. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data* 5, 2 (2017), 153–163.
- [25] Patricia Hill Collins. 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge.
- [26] Sam Corbett-Davies, Emma Pierson, Avi Feller, Sharad Goel, and Aziz Huq. 2017. Algorithmic Decision Making and the Cost of Fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 797–806. <https://doi.org/10.1145/3097983.3098095>
- [27] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. 2009. *Introduction to Algorithms*. MIT Press.
- [28] Sasha Costanza-Chock, Maya Wagoner, Berhan Taye, Caroline Rivas, Chris Schweidler, Georgia Bullen, and the T4SJ Project. 2018. #MoreThanCode: Practitioners reimagine the landscape of technology for justice and equity. (2018). <https://t4sj.co>
- [29] Bo Cowgill. 2018. The Impact of Algorithms on Judicial Discretion: Evidence from Regression Discontinuities. (2018).
- [30] Kimberlé Williams Crenshaw. 1988. Foreword: Towards a Race-Conscious Pedagogy in Legal Education. *National Black Law Journal* 11, 1 (1988), 1–14.
- [31] Kimberlé Williams Crenshaw. 1988. Race, Reform, and Retrenchment: Transformation and Legitimation in Antidiscrimination Law. *Harvard Law Review* 101, 7 (1988), 1331–1387.
- [32] Lorraine Daston and Peter Galison. 2007. *Objectivity*. Zone Books.
- [33] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via Social Media. In *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. AAAI.
- [34] John Dewey. 1938. *Logic: The Theory of Inquiry*. H. Holt and Company.
- [35] Jessa Dickinson, Mark Diaz, Christopher A. Le Dantec, and Sheena Erete. 2019. "The Cavalry Ain't Coming in to Save Us": Supporting Capacities and Relationships Through Civic Tech. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 123:1–123:21. <https://doi.org/10.1145/3359225>
- [36] Paul Dourish. 2004. What we talk about when we talk about context. *Personal Ubiquitous Computing* 8, 1 (2004), 19–30. <https://doi.org/10.1007/s00779-003-0253-8>
- [37] Paul Dourish. 2016. Algorithms and their others: Algorithmic culture in context. *Big Data & Society* 3, 2 (2016). <https://doi.org/10.1177/2053951716665128>
- [38] Anthony Downs. 1962. The law of peak-hour expressway congestion. *Traffic Quarterly* 16, 3 (1962), 393–409.
- [39] Anthony Dunne and Fiona Raby. 2001. *Design Noir: The Secret Life of Electronic Objects*. Springer Science & Business Media.
- [40] Gilles Duranton and Matthew A. Turner. 2011. The Fundamental Law of Road Congestion: Evidence from US Cities. *The American Economic Review* 101, 6 (2011), 2616–2652.
- [41] Andre Esteve, Brett Kuprel, Roberto A. Novoa, Justin Ko, Susan M. Swetter, Helen M. Blau, and Sebastian Thrun. 2017. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542 (2017), 115. <https://doi.org/10.1038/nature21056>
- [42] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press.
- [43] Andrew Guthrie Ferguson. 2017. *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. NYU Press.
- [44] Casey Fiesler, Natalie Garrett, and Nathan Beard. 2020. What Do We Teach When We Teach Tech Ethics? A Syllabi Analysis. In *The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20)*.
- [45] William R. Frey, Desmond U. Patton, Michael B. Gaskell, and Kyle A. McGregor. 2018. Artificial Intelligence and Inclusion: Formerly Gang-Involved Youth as Domain Experts for Analyzing Unstructured Twitter Data. *Social Science Computer Review* (2018), 0894439318788314.
- [46] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Dauméé III, and Kate Crawford. 2018. Datasheets for Datasets. *arXiv preprint arXiv:1803.09010* (2018).
- [47] April Glaser and Will Oremus. 2018. "A Collective Aghastness": Why Silicon Valley workers are demanding their employers stop doing business with the Trump administration. *Slate* (2018). <https://slate.com/technology/2018/06/the-tech-workers-coalition-explains-how-silicon-valley-employees-are-forcing-companies-to-stop-doing-business-with-trump.html>
- [48] Eric Gordon and Stephen Walter. 2016. Meaningful Inefficiencies: Resisting the Logic of Technological Efficiency in the Design of Civic Systems. In *Civic Media: Technology, Design, Practice*, Eric Gordon and Paul Mihailidis (Eds.), 243.
- [49] Andre Gorz. 1967. *Strategy for Labor*. Beacon Press.
- [50] Ben Green. 2018. Data Science as Political Action: Grounding Data Science in a Politics of Justice. *arXiv preprint arXiv:1811.03435* (2018).
- [51] Ben Green. 2018. 'Fair' Risk Assessments: A Precarious Approach for Criminal Justice Reform. In *5th Workshop on Fairness, Accountability, and Transparency in Machine Learning*.
- [52] Ben Green. 2018. Putting the J(ustice) in FAT. *Berkman Klein Center Collection - Medium* (2018). <https://medium.com/berkman-klein-center/putting-the-j-justice-in-fat-28da2b8ea6d>
- [53] Ben Green. 2019. *The Smart Enough City: Putting Technology in Its Place to Reclaim Our Urban Future*. MIT Press.
- [54] Ben Green. 2020. The False Promise of Risk Assessments: Epistemic Reform and the Limits of Fairness. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '20)*. ACM. <https://doi.org/10.1145/3351095.3372869>
- [55] Ben Green and Yiling Chen. 2019. Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*. ACM, 90–99. <https://doi.org/10.1145/3287560.3287563>
- [56] Ben Green and Yiling Chen. 2019. The Principles and Limits of Algorithm-in-the-Loop Decision Making. *Proceedings of the ACM on Human-Computer*

- Interaction* 3, CSCW (2019), 50:1–50:24. <https://doi.org/10.1145/3359152>
- [57] Ben Green, Thibaut Horel, and Andrew V. Papachristos. 2017. Modeling Contagion Through Social Networks to Explain and Predict Gunshot Violence in Chicago, 2006 to 2014. *JAMA Internal Medicine* 177, 3 (2017), 326–333. <https://doi.org/10.1001/jamainternmed.2016.8245>
- [58] Ben Green and Lily Hu. 2018. The Myth in the Methodology: Towards a Re-contextualization of Fairness in Machine Learning. In *Machine Learning: The Debates workshop at the 35th International Conference on Machine Learning*.
- [59] Daniel Greene, Anna Lauren Hoffmann, and Luke Stark. 2019. Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. 2122–2131.
- [60] David Singh Grewal, Amy Kapczynski, and Jedediah Purdy. 2017. Law and Political Economy: Toward a Manifesto. *Law and Political Economy* (2017). <https://lpeblog.org/2017/11/06/law-and-political-economy-toward-a-manifesto/>
- [61] Robert L. Hale. 1923. Coercion and Distribution in a Supposedly Non-Coercive State. *Political Science Quarterly* 38, 3 (1923), 470–494. <https://doi.org/10.2307/2142367>
- [62] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies* 14, 3 (1988), 575–599.
- [63] Sandra Harding. 1998. *Is Science Multicultural?: Postcolonialisms, Feminisms, and Epistemologies*. Indiana University Press.
- [64] Christina Harrington, Sheena Erete, and Anne Marie Piper. 2019. Deconstructing Community-Based Collaborative Design: Towards More Equitable Participatory Design Engagements. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 216:1–216:25. <https://doi.org/10.1145/3359318>
- [65] Nabil Hassenin. 2017. Against Black Inclusion in Facial Recognition. *Digital Talking Drum* (2017). <https://digitalkingdrum.com/2017/08/15/against-black-inclusion-in-facial-recognition/>
- [66] Anna Lauren Hoffmann. 2018. Data Violence and How Bad Engineering Choices Can Damage Society. *Medium* (2018). <https://medium.com/s/story/data-violence-and-how-bad-engineering-choices-can-damage-society-39e44150e1d4>
- [67] Anna Lauren Hoffmann. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society* 22, 7 (2019), 900–915. <https://doi.org/10.1080/1369118X.2019.1573912>
- [68] Wesley Newcomb Hohfeld. 1913. Some Fundamental Legal Conceptions as Applied in Judicial Reasoning. *The Yale Law Journal* 23, 1 (1913), 16–59.
- [69] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. 2018. The dataset nutrition label: A framework to drive higher data quality standards. *arXiv preprint arXiv:1805.03677* (2018).
- [70] Oliver Wendell Holmes. 1897. The Path of the Law. *Harvard Law Review* 10 (1897), 457–478.
- [71] Matthew Hutson. 2018. Artificial intelligence could identify gang crimes—and ignite an ethical firestorm. *Science* (2018). <https://www.sciencemag.org/news/2018/02/artificial-intelligence-could-identify-gang-crimes-and-ignite-ethical-firestorm>
- [72] Armann Ingolfsson, Susan Budge, and Erhan Erkut. 2008. Optimal ambulance location with random delays and travel times. *Health Care Management Science* 11 (2008), 262–274.
- [73] Pretrial Justice Institute. 2017. Pretrial Risk Assessment Can Produce Race-Neutral Results. (2017). <https://university.pretrial.org/HigherLogic/System/DownloadDocumentFile.aspx?DocumentFileKey=5cebc2e7-dfa4-65b2-13cd-300b81a6ad7a>
- [74] C.J. Jagtenberg, Sandjai Bhulai, and R.D. van der Mei. 2017. Optimal Ambulance Dispatching. In *Markov Decision Processes in Practice*. Springer, 269–291.
- [75] Sheila Jasanoff. 2004. Ordering knowledge, ordering society. In *States of Knowledge: The Co-Production of Science and the Social Order*, Sheila Jasanoff (Ed.). Routledge, 13–45.
- [76] Sheila Jasanoff. 2007. Making Order: Law and Science in Action. In *The Handbook of Science and Technology Studies* (third ed.), Edward J. Hackett, Olga Amsterdamska, Michael E. Lynch, and Judy Wajcman (Eds.). MIT Press, 761–786.
- [77] Elizabeth Joh. 2017. The Undue Influence of Surveillance Technology Companies on Policing. *New York University Law Review* (2017).
- [78] Justin Jouvenal. 2016. Police are using software to predict crime. Is it a 'holy grail' or biased against minorities? *The Washington Post* (2016). https://www.washingtonpost.com/local/public-safety/police-are-using-software-to-predict-crime-is-it-a-holy-grail-or-biased-against-minorities/2016/11/17/525a6649-0472-440a-aae1-b283aa8e5de8_story.html
- [79] Becky Kazansky, Guillén Torres, Lonke van der Velden, Kersti Wissenbach, and Stefania Milan. 2019. Data for the Social Good: Toward a Data-Activist Research Agenda. *Good Data* 4 (2019), 244.
- [80] Duncan Kennedy. 2002. The Critique of Rights in Critical Legal Studies. In *Left Legalism/Left Critique*, Wendy Brown and Janet Halley (Eds.). Duke University Press, 178–228.
- [81] Duncan Kennedy. 2006. Three Globalizations of Law and Legal Thought: 1850–2000. In *The New Law and Economic Development: A Critical Appraisal*, David M. Trubek and Alvaro Santos (Eds.). 19–73.
- [82] David Kennedy and William W. Fisher III. 2006. *The Canon of American Legal Thought*. Princeton University Press.
- [83] Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass R. Sunstein. 2019. Discrimination in the Age of Algorithms. *Journal of Legal Analysis* 10 (2019), 113–174. <https://doi.org/10.1093/jla/laz001>
- [84] Jon Kleinberg and Éva Tardos. 2006. *Algorithm Design*. Pearson Education, Inc.
- [85] Adam Kortylewski, Bernhard Egger, Andreas Schneider, Thomas Gerig, Andreas Morel-Forster, and Thomas Vetter. 2019. Analyzing and Reducing the Damage of Dataset Bias to Face Recognition With Synthetic Data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*.
- [86] Pat Langley. 2011. The changing science of machine learning. *Machine Learning* 82, 3 (2011), 275–279. <https://doi.org/10.1007/s10994-011-5242-y>
- [87] Philip Leith. 1990. *Formalism in AI and Computer Science*. Ellis Horwood.
- [88] Lawrence Lessig. 2009. *Code*. Basic Books.
- [89] Empower LLC. 2018. Who's Behind ICE? The Tech and Data Companies Fueling Deportations. *Mijente* (2018). <https://mijente.net/notechforce/>
- [90] Karl N. Llewellyn. 1930. A Realistic Jurisprudence—The Next Step. *Columbia Law Review* 30 (1930), 431.
- [91] Karl N. Llewellyn. 1931. Some Realism about Realism: Responding to Dean Pound. *Harvard Law Review* 44, 8 (1931), 1222–1264.
- [92] Alexandra Mateescu and Madeleine Clare Elish. 2019. AI in Context: The Labor of Integrating New Technologies. *Data & Society* (2019). https://datasociety.net/wp-content/uploads/2019/01/DataandSociety_AlinContext.pdf
- [93] J. Nathan Matias and Merry Mou. 2018. CivilServant: Community-Led Experiments in Platform Governance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, 9:1–9:13. <https://doi.org/10.1145/3173574.3173583>
- [94] Amanda Meng and Carl DiSalvo. 2018. Grassroots resource mobilization through counter-data action. *Big Data & Society* 5, 2 (2018). <https://doi.org/10.1177/2053951718796862>
- [95] Michele Merler, Nalini Ratha, Rogerio S. Feris, and John R. Smith. 2019. Diversity in Faces. *arXiv preprint arXiv:1901.10436* (2019).
- [96] Jacob Metcalf, Emanuel Moss, and danah boyd. 2019. Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics. *Social Research* 86, 2 (2019), 449–476.
- [97] Mijente. 2019. 1,200+ Students at 17 Universities Launch Campaign Targeting Palantir. (2019). <https://notechforce.com/20190916-2/>
- [98] Martha Minow. 1997. The Path as Prologue. *Harvard Law Review* 110, 5 (1997), 1023–1027. <https://doi.org/10.2307/1342112>
- [99] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasseran, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timmit Gebre. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 220–229. <https://doi.org/10.1145/3287560.3287596>
- [100] David Moats and Nick Seaver. 2019. “You Social Scientists Love Mind Games”: Experimenting in the “divide” between data science and critical algorithm studies. *Big Data & Society* 6, 1 (2019), 2053951719833404. <https://doi.org/10.1177/2053951719833404>
- [101] George O. Mohler, Martin B. Short, P. Jeffrey Brantingham, Frederic Paik Schoenberg, and George E. Tita. 2011. Self-exciting point process modeling of crime. *J. Amer. Statist. Assoc.* 106, 493 (2011), 100–108.
- [102] Evgeny Morozov. 2014. *To Save Everything, Click Here: The Folly of Technological Solutionism*. PublicAffairs.
- [103] Paul Mozur. 2019. One Month, 500,000 Face Scans: How China Is Using A.I. to Profile a Minority. *The New York Times* (2019). <https://www.nytimes.com/2019/04/14/technology/china-surveillance-artificial-intelligence-racial-profiling.html>
- [104] Gina Neff, Anissa Tanweer, Brittany Fiore-Gartland, and Laura Osburn. 2017. Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science. *Big Data* 5, 2 (2017), 85–97. <https://doi.org/10.1089/big.2016.0050>
- [105] Design Justice Network. 2016. Design Justice Network Principles. (2016). <http://designjusticenetwork.org/network-principles/>
- [106] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press.
- [107] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- [108] Peter D. Norton. 2011. *Fighting Traffic: The Dawn of the Motor Age in the American City*. MIT Press.
- [109] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 6464 (2019), 447–453. <https://doi.org/10.1126/science.aax2342>
- [110] Cathy O’Neil. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Broadway Books.
- [111] Mimi Onuoha. 2018. Notes on Algorithmic Violence. (2018). <https://github.com/MimiOnuoha/On-Algorithmic-Violence>

- [112] Ginger Adams Otis and Nancy Dillon. 2019. Google using dubious tactics to target people with 'darker skin' in facial recognition project: sources. *New York Daily News* (2019). <https://www.nydailynews.com/news/national/ny-google-darker-skin-tones-facial-recognition-pixel-20191002-5vxpgowknfnvbm5eg7epsf34-story.html>
- [113] George Packer. 2013. Change the World. *The New Yorker* (2013). <http://www.newyorker.com/magazine/2013/05/27/change-the-world>
- [114] Samir Passi and Solon Barocas. 2019. Problem Formulation and Fairness. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 39–48. <https://doi.org/10.1145/3287560.3287567>
- [115] Samir Passi and Steven Jackson. 2017. Data Vision: Learning to See Through Algorithmic Abstraction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 2436–2447. <https://doi.org/10.1145/2998181.2998331>
- [116] Trevor J. Pinch and Wiebe E. Bijker. 1987. The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. In *The Social Construction of Technological Systems*, Wiebe E. Bijker, Thomas P. Hughes, and Trevor Pinch (Eds.). MIT Press.
- [117] Christopher S. Porrino. 2017. Attorney General Law Enforcement Directive 2016-6 v3.0. (2017). https://www.nj.gov/lps/dcj/agguide/directives/ag-directive-2016-6_v3-0.pdf
- [118] Theodore M. Porter. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press.
- [119] Richard A. Posner. 1997. The Path Away from the Law. *Harvard Law Review* 110 (1997), 1039.
- [120] Roscoe Pound. 1909. Liberty of Contract. *Yale Law Journal* 18, 7 (1909).
- [121] Roscoe Pound. 1910. Law in Books and Law in Action. *American Law Review* 44, 1 (1910), 12–36.
- [122] Inioluwa Deborah Raji and Joy Buolamwini. 2019. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, 429–435. <https://doi.org/10.1145/3306618.3314244>
- [123] David Robinson and Logan Koepke. 2016. Stuck in a Pattern. *Upturn* (2016). <https://www.teamupturn.org/reports/2016/stuck-in-a-pattern/>
- [124] Kevin Rose. 2019. The Making of a YouTube Radical. *The New York Times* (2019). <https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html>
- [125] Jathan Sadowski and Roy Bendor. 2019. Selling Smartness: Corporate Narratives and the Smart City as a Sociotechnical Imaginary. *Science, Technology, & Human Values* 44, 3 (2019), 540–563. <https://doi.org/10.1177/0162243918806061>
- [126] Pedro Saleiro, Benedict Kuester, Abby Stevens, Ari Anisfeld, Loren Hinkson, Jesse London, and Rayid Ghani. 2018. Aequitas: A Bias and Fairness Audit Toolkit. *arXiv preprint arXiv:1811.05577* (2018).
- [127] Frederick Schauer. 1987. Formalism. *Yale Law Journal* 97, 4 (1987), 509–548.
- [128] Noam Scheiber and Kate Conger. 2019. Uber and Lyft Drivers Gain Labor Clout, With Help From an App. *The New York Times* (2019). <https://www.nytimes.com/2019/09/20/business/uber-lyft-drivers.html?smid=nytcore-ios-share>
- [129] Bruce Schneier. 2018. *Click Here to Kill Everybody: Security and Survival in a Hyper-connected World*. WW Norton & Company.
- [130] William Schofield. 1907. Christopher Columbus Langdell. *The American Law Register* 55, 5 (1907), 273–296.
- [131] James C. Scott. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.
- [132] Nick Seaver. 2015. The nice thing about context is that everyone has it. *Media, Culture & Society* 37, 7 (2015), 1101–1109. <https://doi.org/10.1177/0163443715594102>
- [133] Nick Seaver. 2017. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* 4, 2 (2017), 2053951717738104. <https://doi.org/10.1177/2053951717738104>
- [134] Nick Seaver. 2019. Knowing Algorithms. In *digitalSTS: A Field Guide for Science & Technology Studies*, Janet Vertesi and David Ribes (Eds.). 412–422.
- [135] Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. Fairness and Abstraction in Sociotechnical Systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. ACM, 59–68. <https://doi.org/10.1145/3287560.3287598>
- [136] Brian Cantwell Smith. 1985. The limits of correctness. *ACM SIGCAS Computers and Society* 14 (1985), 18–26.
- [137] Thomas Smyth and Jill Dimond. 2014. Anti-oppressive design. *Interactions* 21, 6 (2014), 68–71.
- [138] Bryan Stevenson. 2019. Why American Prisons Owe Their Cruelty to Slavery. *The New York Times Magazine* (2019). <https://www.nytimes.com/interactive/2019/08/14/magazine/prison-industrial-complex-slavery-racism.html>
- [139] Eliza Strickland. 2019. How IBM Watson Overpromised and Underdelivered on AI Health Care. *IEEE Spectrum* (2019). <https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care>
- [140] Remi Tachet, Paolo Santi, Stanislav Sobolevsky, Luis Ignacio Reyes-Castro, Emilio Frazzoli, Dirk Helbing, and Carlo Ratti. 2016. Revisiting Street Intersections Using Slot-Based Systems. *PLOS ONE* 11, 3 (2016), e0149607. <https://doi.org/10.1371/journal.pone.0149607>
- [141] Roberto Mangabeira Unger. 1987. *False Necessity: Anti-Necessitarian Social Theory in the Service of Radical Democracy*. Cambridge University Press.
- [142] U.S. Supreme Court. 1905. *Lochner v. New York*. 198 U.S. 45.
- [143] Arnold Ventures. 2019. Statement of Principles on Pretrial Justice and Use of Pretrial Risk Assessment. (2019). <https://craftmediabucket.s3.amazonaws.com/uploads/Arnold-Ventures-Statement-of-Principles-on-Pretrial-Justice.pdf>
- [144] James Vincent. 2016. Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge* (2016). <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>
- [145] Alan Weir. 2015. Formalism in the Philosophy of Mathematics. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/formalism-mathematics/>
- [146] Rebecca Wexler. 2018. Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System. *Stanford Law Review* 70 (2018), 1343–1429.
- [147] Tom Wilson and Madhumita Murgia. 2019. Uganda confirms use of Huawei facial recognition cameras. *Financial Times* (2019). <https://www.ft.com/content/e20580de-c35f-11e9-a8e9-296ca66511c9>
- [148] Jeanette Wing. 2011. Computational Thinking—What and Why? *The Link Magazine* (2011), 20–23.
- [149] Langdon Winner. 1986. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. University of Chicago Press.
- [150] Jimmy Wu. 2019. Optimize What? *Commune* (2019). <https://communemag.com/optimize-what/>
- [151] Bernardo Zacka. 2017. *When the State Meets the Street: Public Service and Moral Agency*. Harvard University Press.
- [152] Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. 2019. Defending Against Neural Fake News. *arXiv preprint arXiv:1905.12616* (2019).
- [153] Lu Zhen, Kai Wang, Hongtao Hu, and Daofang Chang. 2014. A simulation optimization framework for ambulance deployment and relocation problems. *Computers & Industrial Engineering* 72 (2014), 12–23. <https://doi.org/10.1016/j.cie.2014.03.008>
- [154] Eli Zimmerman. 2018. Teachers Are Turning to AI Solutions for Assistance. *EdTech Magazine* (2018). <https://edtechmagazine.com/k12/article/2018/06/teachers-are-turning-ai-solutions-assistance>
- [155] James Zou and Londa Schiebinger. 2018. AI can be sexist and racist – it's time to make it fair. *Nature* 559 (2018), 324–326. <https://www.nature.com/articles/d41586-018-05707-8>
- [156] Mark Zuckerberg. 2018. Protecting democracy is an arms race. Here's how Facebook can help. *The Washington Post* (2018). https://www.washingtonpost.com/opinions/mark-zuckerberg-protecting-democracy-is-an-arms-race-heres-how-facebook-can-help-win-it/2018/09/04/53b3c8ee-b083-11e8-9a6a-565d92a3585d_story.html