

# Evidence of What, for Whom? The Socially Contested Role of Algorithmic Bias in a Predictive Policing Tool

Marta Ziosi\*  
Oxford University  
Oxford, UK  
marta.ziosi@sant.ox.ac.uk

Dasha Pruss\*  
Harvard University  
Cambridge, USA  
dpruss@fas.harvard.edu

## ABSTRACT

This paper presents a critical, qualitative study of the social role of algorithmic bias in the context of the Chicago crime prediction algorithm, a predictive policing tool that forecasts when and where in the city crime is most likely to occur. Through interviews with 18 Chicago-area community organizations, academic researchers, and public sector actors, we show that stakeholders from different groups articulate diverse problem diagnoses of the tool’s algorithmic bias, strategically using it as evidence to advance criminal justice interventions that align with stakeholders’ positionality and political ends. Drawing inspiration from Catherine D’Ignazio’s taxonomy of “refusing and using” data, we find that stakeholders use evidence of algorithmic bias to reform the policies around police patrol allocation; reject algorithm-based policing interventions; reframe crime as a structural rather than interpersonal problem; reveal data on authority figures in an effort to subvert their power; repair and heal families and communities; and, in the case of more powerful actors, to reaffirm their own authority or existing power structures. We identify the implicit assumptions and scope of these varied uses of algorithmic bias as evidence, showing that they require different (and sometimes conflicting) values about policing and AI. This divergence reflects long-standing tensions in the criminal justice reform landscape between the values of liberation and healing often centered by system-impacted communities and the values of surveillance and deterrence often instantiated in data-driven reform measures. We advocate for centering the interests and experiential knowledge of communities impacted by incarceration to ensure that evidence of algorithmic bias can serve as a device to challenge the status quo.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**.

## KEYWORDS

criminal justice; predictive policing; algorithmic bias; qualitative

\*Both authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution International 4.0 License.

FAccT ’24, June 03–06, 2024, Rio de Janeiro, Brazil  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-0450-5/24/06  
<https://doi.org/10.1145/3630106.3658991>

## ACM Reference Format:

Marta Ziosi and Dasha Pruss. 2024. **Evidence of What, for Whom? The Socially Contested Role of Algorithmic Bias in a Predictive Policing Tool**. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT ’24)*, June 03–06, 2024, Rio de Janeiro, Brazil. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3630106.3658991>

## 1 INTRODUCTION

Over the past two decades, public scrutiny and outrage toward policing scandals has led some of the largest police departments in the US to turn to “data-driven” policing as a reform and public relations measure [10]. Place-based predictive policing tools, which forecast crime “hotspots” based on past crime incidence data, have become an increasingly popular and controversial strategy for addressing a range of issues in policing, from resource limitations to unconscious biases in decision-making [63].

Proponents laud predictive policing as an evidence-based policing reform that promises more objective decision-making [18, 40, 55, 56], while critics argue that it exacerbates the very issues it is meant to address [33, 60, 64]. Algorithmic bias – the disparate treatment of or impact on protected groups by an algorithm – has been at the center of the discourse about predictive policing, including concerns about the algorithms’ reliance on “dirty data” and the possibility of feedback loops in officers’ arresting behavior [17, 43, 60, 63].

Standard algorithmic fairness approaches have focused on identifying, removing, or managing statistical biases [21], with strategies ranging from pre-, during and post-deployment strategies to preventing models from learning or outputting biases [17, 20, 79, 80]. However, many critics have pointed out that verifying that an algorithm is correctly ‘calibrated’ is not the right question to ask when underlying data reflect structural injustices [2, 36]. Indeed, the bias exhibited by predictive policing algorithms is considered by many scholars to be an inevitable artifact of higher police presence in historically marginalized communities [43, 81], a reflection of the structural race- and class-biased patterns of the criminal legal system and an unequal society more broadly, rather than a technical “glitch” [8]. In response, some researchers have called for more substantive notions of algorithmic fairness that take algorithms’ objectives and sociotechnical contexts into account [3, 30, 31]. Others have focused on providing affected communities with alternative tools to respond to the deeper societal issues indicated by algorithmic bias, such as the use of civilian dashcam data to detect disparities in police deployments [24]. In short, there has been widespread disagreement about how to respond to evidence of algorithmic bias.

This paper presents a critical, qualitative study on the socially contested role of algorithmic bias in the context of the Chicago

crime prediction algorithm [61], a predictive policing tool that predicts the geographic regions in Chicago in which crime is most likely to occur. The algorithm’s developers demonstrate that it exhibits “enforcement bias” – that is, the uneven or disproportionate patrolling, arrests, or other actions by police toward certain neighborhoods.

Through interviews with Chicago-area academic researchers, community organizations, and public sector employees, we find that evidence of algorithmic bias has the potential to challenge as much as to affirm the status quo. We show that stakeholders from different groups articulate different problem diagnoses of the algorithm’s bias, strategically using these diagnoses to advance criminal justice interventions that align with individuals’ political ends. These strategies include using data to reform the police that generated the biased data in the first place, such as by allocating police patrols more equitably; reaffirm or strengthen existing models of policing; reject the use of predictive policing tools; and reframe the narrative away from reforms like data-driven policing toward abolitionist aims, such as repairing communities.

We identify the implicit assumptions and scope of these varied uses of algorithmic bias as evidence, showing that they require different (and sometimes conflicting) values about policing and AI. Notably, using enforcement bias to inform more efficient allocation of police patrols diverges from the use and interpretation of the same evidence by stakeholders with lived experience in the criminal legal system, who diagnose the source of the bias as systemic. This divergence reflects long-standing tensions in the criminal justice reform landscape between the abolitionist values of liberation and healing often centered by system-impacted communities and the values of surveillance and deterrence often instantiated in data-driven reform measures. We discuss the implications of these differences in relation to participants’ social positions, and we advocate for centering the interests and experiential knowledge of communities impacted by incarceration through participatory methods to ensure that algorithmic bias evidence can function as a device to challenge the status quo.

## 2 BACKGROUND AND RELATED WORK

### 2.1 Algorithmic bias: from an object of social critique to a device for social change

Notwithstanding the widespread claim that algorithmic bias is a mirror of existing disparities [45], it is rarer for algorithmic bias to be taken seriously as a source of evidence to emancipate and support marginalized communities. Bias has helpfully been identified as a boundary object [32], an entity which, despite disagreement about its nature, provides a lens for critique and confrontation of different stakeholders’ views. In this paper, we take this characterization of bias a step further; namely, we want to move from bias as an object of and for critique to a pragmatic device to address the disparities it mirrors.

Skepticism about applying a computational lens to societal problems [8, 11, 14, 19, 26, 29] has led to a reconsideration of the roles of computational tools from optimization and individual predictions toward systemic interventions to address deeper issues of justice and equity [4, 5, 29]. Recent work in critical data studies [1, 14, 54], as well as past work at the intersection of sociology and statistics

[12, 13] has mapped out ways in which practices around the use of computational tools, from the design to the deployment of these tools, can be used to address existing disparities or injustices in the status quo.

Computational tools can help unveil or criticize specific aspects of society. Through their reliance on historical and empirical data, tools both in statistics [12] and in machine learning [1] can help disclose the wrongdoings of an institution by appropriating or questioning its own numbers. The categorization and labeling needed to make data legible can indirectly lead to the creation or resurfacing of neglected or unrecognized social categories, redefine an existing category, or provide a means to defend a category and its rights (e.g., make it “legible” to the system) [12, 14]. The use of computational tools can also reformulate social problems and reshape public objectives. The exercise of problem formulation and formalization the tools entail can help reason about and reassign an institution’s priorities [12], articulate and reason about how societal problems are formulated [54], and highlight the limitations presented by computational techniques [1].

The power of numerical tools, however, can work in both directions. It can expose as much as it can affirm existing power structures [4, 12]. The same numbers that can be used by communities to expose their oppression can, for example, be used by police to justify their intervention. Starting from these considerations, this article explores whether algorithmic bias can be used as a device to disclose or address the societal disparities that it mirrors. At the same time, we explore how evidence of bias enables diverse problem diagnoses and supports the identification of solutions contingent upon individual worldviews. Given its potential to disclose as much as affirm power, we explore how – if at all – this evidence can serve the interests of affected communities.

We connect these questions to a parallel tension in criminal justice between abolitionist and reformist responses to the crisis of mass incarceration. Abolitionist philosophy calls for an end to the reliance on imprisonment and policing and proposes alternatives that engage holistically with the structural conditions underlying violence and suffering [46]. In this vein, abolitionists often advocate for ‘non-reformist reform’, change that is not cordoned by what is possible within the current system but rather change that is premised on what should be made possible long-term given human needs [27]. Non-reformist reform prioritizes decarceration and reinvestment of resources back into communities. This is in contrast to ‘reformist reform’, which limits its objectives to the maintenance and practicality of the current system. Most data science applications in the context of criminal justice are reformist reforms [28]; as we see in the interview results, the predictive policing tool itself is seen as a reformist reform, drawing a sharp line with reform advocates in favor and abolitionists opposed.

### 2.2 The Chicago Crime Prediction Tool

Since the early 2000s, police departments around the US have begun adopting new technologies to forecast crime, often in partnership with private companies like PredPol and Palantir [10]. The Chicago Police Department (CPD) first began using CompStat, an early data-driven geographic forecasting policing program, in 2003 [73]. In 2012, the CPD started using the Strategic Subject List, a

person-based predictive policing tool made by the Illinois Institute of Technology, to generate a “heat list” of likely perpetrators or victims of violent crime [69]. The program was terminated in 2019 following investigations showing that the program was inaccurate, ineffective, and racially biased, among other problems [41, 70]. In 2018, the CPD started using ShotSpotter, a controversial gunshot audio detection technology, to dispatch police to areas with suspected recent gunfire, which led to high-intensity investigatory stops and fatal interactions between the CPD and city residents [65]. In March 2021, ShotSpotter was involved in an incident in which police killed 13-year-old Adam Toledo, intensifying community opposition to the technology. After investigations found that Shotspotter was inaccurate, exacerbated the over-policing of neighborhoods of color, and increased unconstitutional stop and frisk by police [22, 23], city officials announced that they would end their contract with the company by September 2024 [50]. These events all happened in the backdrop of the 2017 consent decree, a court order mandating broad police reform in the city of Chicago.

A recent article on event-level prediction of urban crime in Chicago (herein referred to as the Chicago crime prediction algorithm) [61] is the latest development in this spotted history, though the algorithm is not yet deployed by any police department. Using violent and property crime incidence data from the city of Chicago, Rotaru et al. built a spatiotemporal network that inferred patterns from past event occurrences. This network, known as the Granger network, comprised local estimators forming a communicating network to predict future infractions. Using publicly recorded historical event logs, Rotaru et al. extended their analysis beyond the training sample to events in the subsequent year. See Figure 1 for a model card [49] outlining additional details of the algorithm.

In contrast to the CPD’s wave of failed predictive policing experiments, Rotaru et al. present their novel algorithm as taking a different approach to predictive policing. Their stated aim is not simply to predict hotspots of crime with greater accuracy than commercially available tools but also to “garner deep insight into the nature of the dynamical processes through which policing and crime co-evolve in urban spaces,” including how policing interacts with, modulates and reinforces crime. They emphasize that they see their algorithm as serving the dual role of enhancing state power through criminal surveillance while also providing a means to scrutinize states by allowing researchers to “audit them for enforcement biases” [61]. Indeed, Rotaru et al. write that the Chicago crime prediction algorithm indicates the presence of socioeconomic bias in law enforcement. Modeling the number of individuals arrested during each recorded event allowed the researchers to contrast crime rates with arrest rates, revealing that increases in crime rates correlated with increased arrests in wealthier areas, whereas in low socioeconomic status neighborhoods, arrest rates dropped without a corresponding decrease in crime. This pattern persisted across all the years under analysis. The researchers attributed this bias to resource constraints on police, compounded by discretionary prioritization of wealthier neighborhoods, corroborating suspected enforcement bias toward US suburbs with high socioeconomic status [42, 47]. The authors used this evidence of socioeconomic enforcement bias to suggest a need-based reallocation of police patrols across neighborhoods in Chicago.

**Figure 1: Chicago Crime Prediction Algorithm Model Card**

<ul style="list-style-type: none"> <li>● <b>Model Details</b> <ul style="list-style-type: none"> <li>– Spatio-temporal network inference algorithm: infers patterns of past event occurrences and constructs a communicating network (Granger network) of local estimators to predict future infractions.</li> <li>– Published by academic researchers Rotaru et al. at the University of Chicago in <i>Nature Human Behavior</i> in 2022 [61].</li> </ul> </li> <li>● <b>Intended Use</b> <ul style="list-style-type: none"> <li>– To accurately predict hotspots of crime in Chicago, producing insights on the processes of crime and its interaction with policing.</li> <li>– The authors recommend policy changes for “more equitable, need-based resource allocation” of police and suggest that their tool can “be used to track the extent to which such policies approach this trace of equitable enforcement allocation.”</li> </ul> </li> <li>● <b>Training Data</b> <ul style="list-style-type: none"> <li>– Spatiotemporal logs of crime incidence data in Chicago, as well as the nature, category and description of each incident. <ul style="list-style-type: none"> <li>* Public domain criminal event logs were also analyzed for seven additional cities: Detroit, Philadelphia, Atlanta, Austin, San Francisco, Los Angeles and Portland.</li> </ul> </li> <li>– Socioeconomic variables corresponding to regions of Chicago.</li> <li>– The number of individuals arrested during each recorded event is modelled separately, allowing the investigation of enforcement bias.</li> </ul> </li> <li>● <b>Metrics and Evaluation</b> <ul style="list-style-type: none"> <li>– Area Under Curve (AUC), Predictive Accuracy Index (PAI) (normalized event rate in identified hotspots), Prediction Efficiency Index (PEI).</li> <li>– PAI and PEI results were compared against the best-performing teams in a crime forecast challenge hosted by the National Institute of Justice in 2017.</li> <li>– Enforcement bias: “In wealthier neighbourhoods, the response to elevated crime rates is increased arrests, while arrest rates in disadvantaged neighbourhoods drop but the converse does not occur.”</li> </ul> </li> <li>● <b>Ethical Considerations</b> <ul style="list-style-type: none"> <li>– Potential misuse: the authors acknowledge that these results may be falsely interpreted to justify increased police intervention in higher-crime (often more racially and ethnically diverse) neighborhoods, which could lead to increased surveillance and fatal police interactions in marginalized communities.</li> <li>– Bias in crime reporting data: Rotaru et al. also note that incident logs likely contain “biases arising from disproportionate crime reportage and surveillance [of communities of color],” which “no amount of scrubbing or clever statistical controls can reliably erase.”</li> </ul> </li> </ul>
---

### 3 METHODS

To interrogate the ways in which the enforcement bias captured by the Chicago crime prediction algorithm can serve as evidence, we conducted semi-structured interviews with 18 stakeholders, primarily from the Chicago metropolitan area.

**Recruitment and Demographics.** We conducted interviews with stakeholders from various backgrounds, including academia, community organizations, and the public sector. After an initial mapping of relevant actors through desk research, we selected the rest through referrals from initial respondents. At the time of this writing, the Chicago crime prediction algorithm has not been publicly deployed; snowball sampling allowed us to trace a latent network of relevant actors that we otherwise would not have identified [53]. We continued recruiting and interviewing participants from each

Role	Frequency	%
<b>Community Organizations (C)</b>	7	38.8%
Re-entry Associations	4	22.2%
Safety & Community Initiatives	3	16.6%
<b>Researchers (R)</b>	6	33.4%
Tool Developers	1	5.6%
Academic Researchers	5	27.8%
<b>Public Sector (P)</b>	5	27.8%
Public Officials	1	5.6%
Police	2	11.1%
Contractors & Oversight Bodies <sup>1</sup>	2	11.1%
Total	18	100%

**Table 1: Roles of participant population.**

Exposure to the criminal legal system	Frequency	%
System-impacted individuals	2	11.1%
Working directly with system-impacted individuals	7	38.9%
Employed by the criminal legal system (e.g., police)	2	11.1%
Working with the criminal legal system	3	16.7%
No direct exposure (e.g., researchers)	4	22.2%
Total	18	100%

**Table 2: Participant exposure to the criminal legal system**

background until no new themes or information directly relevant to the project emerged from the interviews [66] or our timeframe for data collection passed. In total, we attempted to recruit 62 participants, achieving a response rate of 29%. We made an effort to select participants with variation [75] across roles (see Table 1) and level of personal exposure to the criminal legal system (see Table 2), ranging from no prior involvement to formerly incarcerated. This variation enabled us to consider the role of individuals’ positionality in the results; we indicate each quoted participant’s background (‘C’ for community organization, ‘R’ for researcher, or ‘P’ for public sector). While eight participants had direct experience or knowledge of the algorithm, several interviewees had no prior familiarity with it; these latter interviews helped inform our understanding of Chicago-area policing more broadly.

Due to confidentiality concerns, recruiting participants from the public sector was challenging. This is a common issue in interviews with participants in senior or official roles – so-called “elite” interviews [59]. We accounted for this by recruiting two additional participants, one who is a contractor for the criminal legal system and another who works for an oversight body.

**Interview Process.** Interviews ranged from 30 minutes to 1 hour, with an average length of 39 minutes. The first author conducted all interviews; two interviews were carried out in person and the rest via video call. Interviews were semi-structured [16], allowing us to focus and dig deeply into an initial set of topics, while also letting participants’ expertise guide the interview and allowing new themes to emerge [62]. Specifically, we asked about participants’ understanding of and attitudes toward algorithmic bias generally and police enforcement bias specifically; the challenges and problems they believe to be associated with these biases; and the interventions available or those they envision in the future (see Appendix A for a full list of interview questions). We adjusted the content and

wording of the questions slightly depending on the participants’ particular knowledge of or experience with the algorithm and the criminal legal system. For example, for participants with intimate knowledge of the algorithm, such as its developers, more technically detailed questions were posed, while for participants with local knowledge of the criminal legal system, such as members of community organizations, a stronger emphasis was placed on questions about the Chicago context and community safety.

**Qualitative Analysis.** We analyzed interview transcripts using the qualitative data analysis software Nvivo. Both authors analyzed interviews iteratively, identifying recurring themes, reviewing transcriptions 2-3 times, and switching between inductive coding and data collection to clarify and strengthen themes and uncover opposing evidence [48]. We converged on seven high-level themes (in order of most to least prevalence in the interview data): downsides of AI; upsides of AI; structural bias; lived experience and positionality; abolition; reformism; and controversy around policing (see Appendix B for a code table). We complemented the results from the interviews with a document analysis [39] of policy proposals and programs produced by the interviewed stakeholders, as well as other active participants in the domain of policing and justice reform in the Chicago area (e.g., policymakers, regulators, community organizations, and think tanks). Through triangulation with findings from our interview data, we were able to elucidate, refute, or expand on the findings from our interview data [25].

## 4 RESULTS

Across the board, interviewed stakeholders strategically used algorithmic bias as evidence to advance interventions that aligned with their own positionality and values about the criminal legal system. Drawing inspiration from Catherine D’Ignazio’s taxonomy of femicide activists “refusing and using” data [14], we see that stakeholders from different backgrounds use evidence of algorithmic bias to **reform** the policies around police patrol allocation; in the case of more powerful actors, **reaffirm** their own authority or existing power structures; **reject** algorithm-based policing interventions; **reveal** counterdata on authority figures in an effort to subvert their power; **reframe** crime as a structural rather than interpersonal problem; and **repair** and heal families and communities.

### 4.1 Reform: Towards “a legitimate, ethical, fair police force.” [P013]

**4.1.1 Policing as an optimization problem.** Making policing more ‘fair’ is a standard goal of criminal justice reform [72]. Stakeholders from academic and public sector contexts tended to focus on ‘over-’ or ‘under-policing’ as the source of the Chicago crime prediction algorithm’s bias, emphasizing the benefits of allocating policing more optimally, such as increased public safety and police legitimacy. However, these participants had mixed views about whether predictive policing was a viable strategy to achieve this goal.

Three of the academic researchers we interviewed, including one of the tool’s developers and members of a Chicago-area research group focused on criminal justice research, framed the problem revealed by algorithmic bias as a police allocation issue. One senior member of the research group [R006] explained the need for improving police presence in marginalized neighborhoods: “People

who live in those communities ask for more police because they're facing high levels [of violent crime] ... I mean, the reality is like somehow we are both, like, those communities are both over- and under-policed at the same time." Another member of the research group [R007] posited that "police presence alone deters crime," even if police make no arrests.

These attitudes aligned with perspectives voiced by two police officers. Echoing the senior academic's perspective, one officer [P003] argued that "the residents in those [Black and Brown] neighborhoods, that's the silent majority. They want us there, they need us." Another officer [P013] distinguished "feeling safe" and "being safe," emphasizing the significance of the former and highlighting that police presence contributes to a sense of security among citizens. On this shared view, there is intrinsic value to policing, and distributing it optimally is a primary reform goal.

However, some stakeholders who were otherwise proponents of reforming police allocation expressed reservations about the use of predictive policing specifically to achieve those goals. The officer who argued that increased police presence promotes the feeling of safety [P013] warned that the use of predictive policing and other technological systems could be "preventing cops from doing their job, which is a social job" by making them too "comfortable coming into an office and sitting behind a computer all day," thereby diminishing the physical presence of police in communities and compromising the feeling of safety. Likewise, one of the tool's developers [R002] did not want the tool to be directly used to inform police allocation decisions.

**4.1.2 Reducing interpersonal bias.** Several public sector actors saw the tool's enforcement bias as a potential result of another kind of unfairness – interpersonal bias, including implicit bias, the bad judgment of a "few bad apples," or simply limitations in human cognition. Reducing interpersonal bias among police officers through implicit bias training, data-driven policing, or more diverse police forces is another standard route of reform [67, 72].

One county commissioner [P015] in a majority Latino district of Chicago called implicit bias "very damaging" to communities and explained that police officers, judges, and "all of our criminal justice system" must attend implicit bias training. A police officer [P013] also described what he called declining "standards of professional behavior" among police officers, underscoring the need for strong leadership programs, adding that this would improve community trust and legitimacy of police.

Arguing for the benefits of "research-driven" policing, another police officer [P003] admitted that "we're not the brightest bulbs" but that "there are smarter minds analyzing all the different mathematical models in the back end of the AI, so they'll be able to inform us and they'll keep all of the AI models accountable." Using the example of people's tendency to overcount the significance of a stale criminal record, one academic researcher [R006] argued that an algorithm could "have a sort of more built-in decay mechanism than like a human could," thus improving on limitations in human cognition.

Participants across all roles also proposed diversity, equity and inclusion (DEI) as a strategy for minimizing interpersonal bias, especially integrating diverse perspective or other stakeholders in the creation of data and algorithms. As one interviewee from a

reentry association [C001] put it, "whose voice should be in the room for the algorithms that are being built, right? And I don't just mean a little bit of, yeah, engaging in participatory research where you sit down with a group of folks and this and that and you go off in your room and you build this thing. No, I mean you actually have to make them or the folks from these communities where these tools are going to be deployed an active creator of the system, right?" [C001] Several researchers working with system-impacted individuals and victims of violence also stressed the importance of incorporating those groups' voices in research.

Another participant from a reentry association [C004], however, highlighted that incorporating diverse stakeholders on its own would not solve the problem: "You could take a Black person or a Brown person, or some other disenfranchised individual, and they could be the author of, you know, some new algorithm. But unless they have a specific training and knowledge around systems of bias – how they're constructed, how they operate and how they can be deconstructed – then I don't think that these individuals that are constructing these programs can, can effectively create solutions for them."

**4.1.3 Increasing transparency and accountability.** Beyond algorithmic bias, some participants associated another reform goal with the predictive policing tool, namely, accountability and transparency. This is a common byline in police reform rhetoric [76] and is often presented as a justification for data-driven policing and surveillance through body-worn cameras.

Referring to the Chicago crime prediction tool, one police officer [P003] said that "it being open source is a winner" because of what he pejoratively referred to as the "two big buzzwords" around policing following George Floyd's murder in 2020, namely, accountability and transparency. "Here's the open source data, you know. You look at it then I'm not hiding anything. There's no secret sauce to this. There's no ulterior motives."

However, other participants pointed out that the relationship between open-source and accountability is far from a given. One reentry community organizer [C001] argued that "the people that build them [algorithms] should have some accountability for how they're used ... even if you made it open source, right? Even more then, in some ways, right? You have an obligation to have thought about how it could be used for better or for worse, right?" One of the tool's developers [R002] acknowledged the same issue with open-source, noting that "you have to also be careful when you're giving it away" because nothing prevents private companies from forking open-source software into a private (i.e., non-transparent) repository. The fact that police officers explicitly reported wanting to acquire the predictive policing tool underscores that even open-source software, once developed, may be used in ways contrary to the desires of their developers, particularly once it becomes commercialized [77, 78].

## 4.2 Reaffirm: "How can there be a bias if in that triangle it's all the same race?" [P003]

**4.2.1 Following the data.** Participants in public sector positions denied the presence of bias or used evidence of algorithmic bias to reaffirm their own authority and the legitimacy of existing policing

practices. This is exemplified well by one officer [P003], who candidly explained why there is no over-policing problem: “Where the crime is happening, the concentration is 99% Black or Brown and the victims are 99% Black or Brown, so that’s our baseline. ... We’re there because the data says that’s where we need to be.” He added that “it’s not that the AI is biased. The AI is just working with the data, and the data says that this geography, this demographic is, you know, these individuals.”

This officer provided a problem diagnosis of the algorithm’s behavior that essentialized the policed community – “these individuals” – as inherently exhibiting more criminal behavior, thereby using the pattern to justify or take as given the status quo. Disparities in crime, in other words, are the factor that drives disparities in policing and in turn the algorithmic non-bias, and the officer lamented that “in policing we’ve been knocked around” for simply responding to what “the data is telling us.”

**4.2.2 New and improved technology.** Another response to algorithmic bias that reaffirms police power is the call to invest in more technology. Some public sector participants saw the algorithm as an opportunity to do their work better, and others as the only way forward. For instance, speaking about the motivation behind the development of the crime prediction tool, one developer explained that “all of them [previous predictive policing algorithms] were doing it wrong” and the need to “show what is the right way to do it.” One police officer [P003] called for the need for policing to be informed by the crime prediction algorithm and to “try a different approach on the AI.”

**4.2.3 More policing.** As we discuss in §4.1.1, one response to enforcement bias is the call to increase police presence in areas with higher crime incidence, which implicitly affirms the legitimacy of the current policing system. For instance, one police officer reported being interested in “using this technology to break down the cycles that have been perpetuated for years” [P003]. However, it became clear that what they meant by this was sending more police to Black and Brown neighbourhoods in order to deter crime, rather than to change the structure of the current system or “break down” any cycles.

An interviewee from a community association [C011] voiced concern about affected communities similarly endorsing approaches that do not change the existing system and go against their own interests, such as “more policing,” and attributed this to a lack of awareness of alternative options. Referring to the race for a major election in a large US city, they said, “... older Black residents who are concerned about public safety, and many of them, their response is ‘we need more police’. And the reason is because it’s the only tool they know, right?”

### 4.3 Reject: “I don’t think it should be used in any way, shape or form.” [C005]

**4.3.1 Garbage in, garbage out.** Stakeholders across all groups shared some inherent reservations about AI in the policing context, particularly concerns about its underlying data. In particular, the theme of algorithmic bias being a “result of the system” [C001] – that is, reflecting structural bias and systemic racism – recurred frequently. One system-impacted participant [C005] argued that the

system should not be used at all because it does not account for systemic and individual bias: “It just cannot be neutral, and therefore it should not be used because people that look like me are the ones who are going to be impacted.”

Another concern voiced by researchers working with system-impacted individuals was a general distrust of low-quality or manipulated police data. “Their [police] incentives are wrapped up in what is reported. You know what crimes are in the news and what people are saying. It’s like they’re kind of incentivized to obscure data.” [R008] Researcher participants described that documentation of police misconduct is likely to be reduced to the most minor violation; for instance, an illegal search will only be reported if it also includes sexual assault. Another researcher working with system-impacted individuals [R009] pointed to several cases where police officers intentionally misreported the race of individuals in order to conceal bias and discrimination in their practices. He mentioned that in Louisiana, for example, several law enforcement agencies targeted Latino individuals but identified them as white on tickets in order to conceal discrimination [74], and similar incidents have also taken place in Connecticut [52], Los Angeles and San Francisco [6].

Finally, several participants from community organizations had reservations about the inherently interpretative component of AI, which they said biases its results. A member from a reentry community organization [C001] argued that AI is an “interpretive instrument” rather than a mechanical one. “In an ideal sense, let’s say the [AI] system had no politics, right? But the people that will use the results and interpret the information do, and you know, as they, you know, they always say, you know, people can make statistics tell any story they want it to tell.”

**4.3.2 Unneeded, unconvincing, and resource-intensive.** Other respondents were opposed to the predictive policing tool for more pragmatic reasons. For instance, several participants felt that the algorithm’s enforcement bias did not show anything that was not already known. One academic researcher [R006] commented that developing an algorithm is “a pretty inefficient way” to inform police reallocation, arguing that “you wouldn’t need that, to develop an algorithm to reach these same conclusions.” An interviewee working for a reentry association [C017] expressed similar skepticism about the need for a tool that predicts crime. “He [system-impacted individual] is always going to be discriminated against in one form or fashion, maybe an employment house. He’s always going to be there, and that’s a predictor of recidivism or returning to be a criminal. Well, yeah. We don’t need an algorithm to say that.”

The academic researcher [R006] also said that the algorithm was “not particularly convincing” for other stakeholders: “I barely understand exactly what they’re doing and how they reach their conclusion, ... And like, I have a PhD and I do this for a living. So I think, you know, like trying to convince like, other stakeholders is very difficult.” Other participants, including another academic researcher and several community organizations, also reported not understanding the paper or how the tool worked, though in the case of one police officer [P003], that was a positive feature because it indicated that the tool’s developers were “smarter minds.”

The same police officer [P003] also commented on the resource-intensiveness, both in terms of computational power and departmental funding, of acquiring and using a predictive policing tool in practice. “Will we have a budget for this? Or do I need to go out and fundraise for it and convince foundations that, you know, this is the way we need to use the limited resources that we have?”

Understanding the underlying motivation for developing the tool helps contextualize the above criticisms. One of the tool’s developers [R002] explained that the team had been working on a general algorithm and were looking for a real-world test case to demonstrate how the algorithm could be instantiated. “Why crime? No particular reason. We were looking for a problem that would be an impactful thing.” An advantage of working on crime, they add, is that there is “lots of data on it.” The research team had access to compute power that a typical community organization or police force does not, and it is clear that the team did not have the goal of convincing stakeholders or revealing surprising patterns. Rather, the research team saw predicting crime as an “interesting” challenge, seeking the ‘technical sweetness’ [35] that engineers feel when they encounter the solution to a challenging technical puzzle.

#### 4.4 Reveal: “Data that helps us understand how the police see our communities.” [R009]

Enforcement bias was one type of evidence community organizations used to engage in *sousveillance*, or ‘watching from below’, in which the traditional framing of panoptic state surveillance is inverted, allowing citizens to observe powerful state actors [44, 51]. Researchers and community organizers described how they co-opted surveillance methods to watch the watchers. Alternate data sources, or counterdata [14], reveal issues with policing, remediate problems with how police and victim data are collected and reported, and allow alternate narratives to emerge.

One member of a community safety initiative [C101] lamented that “a lot of times what criminologists will tend to do is they’ll just look at the data. And I think what’s missing is that there’s a story behind it.” To provide this missing context, one human rights researcher [R009] described working with grassroots groups that do on-the-ground documentation through cop watching and court watching, gathering counterdata on police officers, judges, and prosecutors. They argue that “you need to record different things than what the courts are” in order to include context normally absent in court data. The same researcher also described using data about human rights violations to understand patterns of violence, and even to make statistical estimates of events that are not represented in the data.

As we discuss in §4.3.1, several researchers working with system-impacted individuals also reported the existence of low-quality, unreliable or manipulated police data. By collecting data from the Chicago Police Department, the city of Chicago, and the Civilian Office of Police Accountability through Freedom of Information requests, one research group [R008] described how they published an alternative dataset about police misconduct in Chicago. They also recounted collaborating with system-impacted individuals and victims of police violence to build a basic classifier to predict police misconduct using complaint data generated by community members. This alternate algorithm revealed a culture of abuse within

policing: “instead of something like operation and personnel violation, you know, we’re able to see clusters of, you know, families who are like ... ‘this is how police treated me’. And what was really fascinating about that project is that we were able to, you know, move away from this idea of like, individual officers, like, back up to what is maybe the culture of policing that exists, what are the kind of tactics that, that people are reporting that connect them across space and time?”

#### 4.5 Reframe and Repair: “Deconstructing the status quo and recreating a new status quo.” [C004]

Countering mainstream understandings of crime among respondents from the public sector and some academic researchers, many participants from community organizations presented alternate data epistemologies [58] focused on abolitionist understandings of the “root of the problem” [C016]. In this vein, participants responded to enforcement bias by calling for the need to recognize structural racism, reframe the narrative toward an understanding of crime as structural rather than interpersonal, and repair and build healthy communities.

*4.5.1 Recognition of systemic racism.* A common thread in the interview data was the “tale of two cities,” [C001; P015] or Chicago’s history of segregation and systematic disinvestment in its Black and Brown communities. Participants emphasized the significance of Chicago’s history of segregation, citing it as a root cause of not only crime but also numerous other social problems, such as disparities in life expectancy – “I want to say 20-30 year difference in life span between those that live in the poor neighbour community” [C001] – and in wealth – “You can go to the different parts of the north side and you can just see the wealth, right? ... and then you can go to the south side or the southwest side and see completely different – you, you can almost think that you’re in a different city” [P015].

Some system-impacted individuals and researchers working with them also remarked on the intrinsic suspicion they experience toward policed communities, accompanied by a tendency to trust or believe those doing the policing. One of them [C017] argued that criminal legal policies are “setting people up for failure,” using Nixon’s 1970s policy on potential delinquents as an example of how ‘criminals’ are created out of assumptions rather than facts.

Multiple participants stressed that the first step to resolving the problem presented by enforcement bias is having a clear understanding of what the problem is. One member of a community safety initiative in Chicago [C011] noted that “before anybody can heal, there needs to be a reckoning, there needs to be some kind of identification that, that the systems that were created actually have caused harm.” Two more participants [C004; C005] echoed this sentiment almost verbatim, arguing that “if we’re not going to address a system of white supremacy, then we’re not going to get to any real hard solutions.” [C004]

*4.5.2 Narrative change.* Beyond reckoning, participants emphasized reframing the narrative around system-impacted individuals, police, and community safety. As one community worker [C005] put it, “The delusion that somehow we can scare people out of crimes of poverty is ridiculous. We have been so indoctrinated by

current systems to accept the claims, regardless of the the reality, right?” They argued that “One of the biggest things that we do is narrative change work” in order to push back on dominant narratives around crime and policing.

In particular, participants call for a reorientation toward a structural understanding of the “root cause” [C014] of crime, focusing on treating the cause rather than the symptom. Two participants criticized interventions that focus on the interpersonal level, with one participant [C011] comparing the current criminal legal system to a hospital that only contains an emergency room and another [R104] lamenting that interventions like community violence and prevention programs “stop at the interpersonal level, but they don’t also account for the structural systemic impediments at play.”

*4.5.3 “Building healthy communities rather than reducing crime” [C001].* Grounded in an alternate understanding of crime as structural, participants responded to enforcement bias by advocating for non-reformist reforms such as alternatives to police and rethinking how to use data to promote community safety.

One community organizer [C001] asked, “instead of thinking about how we reduce violence and crime, what if we pose the question of how this information helps us build healthy communities?” We saw in §4.4 how researchers and activists gather counterdata on police and make up for limitations in data. This strategy centers the needs of affected communities, which community organization participants agreed got little benefit from predicting crime. Another researcher [R008] explained how AI could instead be used to reduce crime in the sense of supporting survivors of violence, especially police violence: “When we pivot our approach to, you know, engaging with survivors, then we’re better understanding the implications of policing and the ways that we can use artificial intelligence to support survivors of violence and reduce crime, like, period.”

Beyond the use of data, several participants also advocated for non-reformist reforms. Drawing on an abolitionist critique of police, one researcher [R008] working with system-impacted individuals challenged the assumptions underlying policing: “Policing should not be about ‘we gotta get the bad guys’, right? It should be about ‘how do we keep people safe.’” Taking this view a step further, one community organizer [C101] questioned the belief that “we need more [police],” adding that “if we bring more tools to the community, there’s— the solutions are here. We just haven’t given you all the equipment to do so.” Notably, the police and academic position on police promoting safety was starkly different from the perspective of some interviewees from community organizations for which police presence was an “enforcement presence, not a deterrent, but an enforcement” [C001].

Even when there was agreement about the need to strengthen communities, however, respondents differed in their views on which kinds of interventions to put in place to do so. Some reentry organizations advocated for giving the same rights to system-impacted individuals or clearing their past criminal record after some time [C017], while others fought for more police and state accountability [R008]. One individual working as a contractor for the public sector [P016] recognized the structural nature of the problem and yet focused on “changing the individual” as a solution, namely, moving individuals out of a high-crime environment or, alternatively,

strategically reclaiming high-crime spaces by organizing community events (e.g., sports competitions, community barbecues, etc.). Another member of a community organization [C011] focused on “changing the conditions” – namely, focusing interventions on the socioeconomic determinants of crime such as wealth and health. In short, even within the non-reformist reform branch of responses to enforcement bias, there were a diversity of perspectives on the appropriate kinds of interventions.

## 5 DISCUSSION

Our results shed light on the socially contested nature of algorithmic bias within the realm of predictive policing, illustrating that bias is not understood as a singular challenge but rather evidence of a multifaceted package of problems with a corresponding broad range of possible interventions. We grouped the interview findings according to whether stakeholders from different backgrounds used evidence of algorithmic bias to reform, reaffirm, reject, reveal, reframe or repair, showing the ability of enforcement bias to be used both to support and challenge the status quo.

Whether and how algorithmic bias was seen as useful for advancing each of these goals varied according to each stakeholder’s broader understanding of the source of police enforcement disparities – structural inequality versus interpersonal bias – and the set of interventions aligned with each respective problem diagnosis. Stakeholders from different professional roles and positionalities tended to articulate problem diagnoses of the bias aligned with their own relationship to the criminal legal system, though we saw a diversity of attitudes and responses within each subgroup and points of convergence between individuals with different positionalities and roles. For instance, we saw substantial overlap among interviewees from all backgrounds regarding the importance of community safety and potential downsides to the use of AI and data in the context of policing, such as AI’s possibility to diminish the role of nuanced human discretion. On the other hand, we saw disagreement between groups about whether police presence promotes community safety, and there were differences within groups relating to whether predictive policing tools would improve policing and whether incorporating affected stakeholders in the development of predictive policing could improve the tools’ viability.

Many of the inter-group differences we saw reflect tensions between the assumptions underlying interventions grounded in reform versus abolition – reformist reform that operates within the logic of the current system and non-reformist reform that aims for longer-term systemic change, respectively. Along these lines, we saw other contested concepts emerge, such as the notion of community safety and ‘overpoliced’ communities, with the connotation and valence of ‘overpoliced’ changing depending on the socioeconomic status of the overpoliced community. For example, in their paper, developers of the Chicago crime prediction algorithm suggest that police are significantly more present in wealthier neighborhoods because those areas have more resources. ‘Overpoliced’ here has a positive connotation – the privilege and safety of well-resourced communities. At the same time, interviewees from community organizations and from the public sector tended to refer to overpoliced communities as poorer communities with higher rates of crime.



While inter-group differences reflect historical tensions between reform and abolition, one additional significant binary emerged that correlated with participants' attitudes toward the criminal legal system – 'AI hype' rhetoric. We examine this finding below and we end by discussing the relationship between participant positionality and research on algorithmic bias in predictive policing.

## 5.1 AI hype rhetoric

With stakeholders from all groups, except the academic researchers most closely familiar with the tool's development, attitudes toward predictive policing tended to fall into one of two AI hype narratives: optimism about AI's potential to 'disrupt' the criminal legal system in a beneficial way, and inherent opposition to AI on the basis that it is racist and fundamentally flawed. Techno-optimism mapped onto the extent to which a stakeholder understood the problems the algorithm was capturing as primarily interpersonal rather than structural, which correlated with positionalities in higher positions of power. However, some endorsers of structural inequality narratives still saw value in certain applications of AI, such as grassroots classifiers to counter narratives put forward by police or to reveal patterns of police violence.

The clearest example of disruption rhetoric came from officer P003, who praised the "smarter minds" developing the predictive policing tool (§4.1.2) and argued that more technology – or trying "a different approach on the AI" [P003] – could help tackle the bias in the algorithm (§4.2.2). They also remarked repeatedly that AI could help the police force's "marketing ability to brand us as open as transparent and as accountable," what they called the two salient "buzzwords" of the "2020, BLM, post-Floyd" era (§4.1.3).

AI may yet become a successful marketing tool for police forces, similarly to past cases in which police acquisitions of technologies like body-worn cameras strategically co-opted community calls for accountability to justify larger police budgets [38, 71].<sup>2</sup> However, the underlying claim that predictive policing promotes transparency and accountability in policing is based in fiction. Sociologist Sarah Brayne provides a compelling argument that predictive policing simply displaces discretion to other, less publicly visible stages of the policing process, thereby reducing both transparency and accountability [10].

On the other end of the AI hype spectrum, we saw community organizers describe AI as inherently racist. As we discussed in §4.3.1, a system-impacted leader of a reentry organization [C005] argued that predictive algorithms should not be used at all in the context of the criminal legal system because "it's just racist and it needs to be stopped." Those in positions of power, they argued, are incentivized to use algorithmic systems because "it keeps the status quo in place," which is what gives those individuals power. An interviewee running a reentry association [C017] echoed that these systems "set people up for failure" (§4.5.1) and that people with a criminal record are "always going to be discriminated against in one form or fashion" (§4.3.2). Like P003, C005's and C017's perspective paints a broad and uniform brush across all uses of AI in the context of the criminal legal system. Both groups at the extreme ends of the spectrum had direct experience with the police but from starkly

different positions of power; from being a police officer to having been previously incarcerated.

Not all the participants we interviewed fell into this AI hype dichotomy, but the polarity was widespread enough that we believe it is important to explicitly take into account in future work that aims to use algorithmic bias evidence to challenge the status quo. In practice, existing empirical research of evidence-based reforms shows that, despite their hype and high price tag, they often have minimal effects that wash out over time, suggesting that these reforms are indeed harmful but not in the ways that critics may expect [57, 68]. As we explain in the following section, participatory research is crucial to develop responses to algorithmic bias that are attentive to the interests and knowledge of communities impacted by incarceration; at the same time, such work would be made more robust by helping participants resist one-dimensional narratives about the harms of technology in the criminal legal system.

## 5.2 "Bias as defined by who, right? So ultimately, who's the judge?" [C001]

The range of problem diagnoses and uses of algorithmic bias might be interpreted as a 'relativistic' stance toward the issue, where the understanding of algorithmic bias is contingent upon an individual's subjective world view. Rather than dismissing these differences, we believe that they provide a valuable lens to understand the conditions under which algorithmic bias becomes a device to reinforce or challenge power dynamics. Drawing on standpoint theory [34], we argue that unraveling the relationship between bias and positionality serves as a critical tool to evaluate shortcomings and mistaken assumptions in potential responses to algorithmic bias coming from actors in relatively high positions of power, such as police officers and even academic researchers such as ourselves.

Standpoint theory posits that individuals' social positions shape their access to knowledge, influencing how they interpret and respond to phenomena such as bias. Social locations 'at the margins' that have histories of repression and resistance have better awareness of injustices and can more easily question the 'naturalness' of existing power structures [9, 15, 37]. In the context of algorithmic bias and predictive policing, this means that those in positions of power, such as public sector actors and academics, may inadvertently reaffirm fundamentally unjust systems due to their distance from the lived experience of incarceration.

Our findings emphasize the importance of centering the needs and experiential understanding of system-impacted communities through community-based participatory research in designing and evaluating responses to algorithmic bias. A critical consideration of positionality and bias may promote interventions that can serve affected communities' stated goals of community investment and healing, contributing to the development of interventions that address the root cause of algorithmic bias in predictive policing. Our results suggest that such a community-centered response to bias would in fact leave no room for the existence of crime prediction tools like predictive policing, which take the current system as a given, and would prompt a reimagining of ethical data science grounded in abolitionist philosophy and a refusal of oppressive structures [4, 7].

<sup>2</sup>Police are trained to manipulate and withhold body-worn camera footage to their own benefit [71].

## 6 CONCLUSION

In this article, we presented a critical, qualitative study on evidence of enforcement bias within the Chicago crime prediction algorithm [61]. Our research, based on interviews with various stakeholders, including community organizations, academic researchers, and public sector figures in the Chicago area, reveals that stakeholders offer diverse diagnoses of this bias, strategically leveraging these diagnoses to promote interventions that align with their respective positionality.

These interventions encompass a range of approaches, such as using data to reform police practices that contributed to biased data collection; advocating for increased transparency and accountability in the development and utilization of predictive policing tools; discontinuing the use of such tools altogether; involving a diverse set of stakeholders in the creation and evaluation of these tools; and proposing a shift away from investing in data-driven policing reforms toward more abolitionist goals, such as community investment. We examined the implicit assumptions and the scope underlying these diverse responses to algorithmic bias, demonstrating that they are grounded in reformist and abolitionist thinking. We concluded by reflecting on the divergent AI hype views reported by participants and its relationship to lived experience with the criminal legal system. We advocate for centering the interests and experiential knowledge of system-impacted communities to ensure that evidence of algorithmic bias can be mobilized as a device to challenge the status quo.

## 7 ETHICAL CONSIDERATIONS

### 7.1 Researcher Positionality

As white women academic researchers from Europe and the US, we acknowledge that our personal and professional backgrounds shaped how we selected research questions and analyzed data, as well as the manner in which we interacted with study participants from backgrounds similar and dissimilar to ours. Both authors have a privileged academic affiliation, race, and lack of personal contact with the US criminal legal system, which undoubtedly shaped the information that grassroots community organizations and system-impacted interviewees were comfortable sharing with us. In fact, the theme of exploitation of marginalized communities by academic researchers came up explicitly in one interview, which we discuss in §7.2. On the other hand, our privileged position as white academic researchers from well-recognized institutions helped build rapport with other academic researchers in the study, who had similarly received training at elite academic institutions.

### 7.2 Ethical Concerns

**Privacy and anonymity.** In line with ACM Ethics Guidelines and the requirements of our IRB approval, participants' identities were protected by pseudonymizing interview transcripts. The document matching pseudonym to identifiable information was stored separately from the research data.

**Engaging non-extractively with affected communities.** The primary ethical challenge we encountered had to do with how to involve stakeholders impacted by the criminal legal system in our research in a non-extractive and mutually beneficial manner.

Due to our own resource constraints, we did not compensate interviewees for their time, which meant there was no direct benefit for study participation, and our positionality meant that affected stakeholders had little incentive or reason to place their trust in us. This became particularly salient at the end of an interview with a system-impacted participant, who works with returning citizens in Chicago. In response to a question about referring us to other study participants, they responded: "I'm going to be completely transparent with you. Our community is exploited, often by academia. And, and so I don't feel comfortable because of that constant exploitation to put other people in that position. ... [Academic researchers] take our expertise, they subject us to trauma by asking us to tell our stories, and then they ... never come back and connect it to us."

We acknowledge the respondent's concern and plan to share all research outputs with research participants, crediting them when desired, with the understanding that we cannot give back as much as we have received from their participation. The experience of this power differential demonstrated to us the importance of building longer-lasting community relationships outside the scope of academic projects prior to engaging in research, and increasing communities' ownership and participation of the research, with the gold standard in mind of participatory action research.

### 7.3 Adverse Impacts

By focusing on the Chicago crime prediction algorithm, our research may inadvertently publicize the availability of the tool or promote more intensive policing. Beyond recognizing this limitation, we wish to make explicit that we do not support the development of this (or similar tools) in any form. We hope that, by shining light on the tool and the different concerns it raises, we can show how and why solutions addressed at tackling bias often do not end up benefiting the affected communities.

## ACKNOWLEDGMENTS

We thank the respondents from community organizations in Chicago and all other interview participants for their time, trust and invaluable knowledge. The research was supported by the first author's affiliation with the Oxford Internet Institute's Programme on AI, Government and Policy, funded by the Dieter Schwarz Stiftung gGmbH, and the second author's fellowship at the Berkman Klein Center for Internet & Society at Harvard University. We thank the Embedded EthiCS group at Harvard University for their feedback on this work.

## REFERENCES

- [1] Rediet Abebe, Solon Barocas, Jon Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. Roles for Computing in Social Change. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 252–260. <https://doi.org/10.1145/3351095.3372871> arXiv:1912.04883 [cs].
- [2] Agathe Balayn and Seda Gürses. 2021. Beyond Debiasing: Regulating AI and its inequalities. *EDRi Report*. [https://edri.org/wp-content/uploads/2021/09/EDRi\\_Beyond-Debiasing-Report-Online.pdf](https://edri.org/wp-content/uploads/2021/09/EDRi_Beyond-Debiasing-Report-Online.pdf) (2021).
- [3] Chelsea Barabas. 2019. *Beyond Bias: Re-Imagining the Terms of 'Ethical AI' in Criminal Law*. SSRN Scholarly Paper ID 3377921. Social Science Research Network, Rochester, NY. <https://doi.org/10.2139/ssrn.3377921>
- [4] Chelsea Barabas. 2022. Refusal in Data Ethics: Re-Imagining the Code Beneath the Code of Computation in the Carceral State. *Engaging Science, Technology, and Society* 8, 2 (Sept. 2022), 35–57. <https://doi.org/10.17351/ests2022.1233> Number: 2.

- [5] Chelsea Barabas, Madars Virza, Karthik Dinakar, Joichi Ito, and Jonathan Zittrain. 2018. Interventions over Predictions: Reframing the Ethical Debate for Actuarial Risk Assessment. In *Conference on Fairness, Accountability and Transparency*. 62–76. <http://proceedings.mlr.press/v81/barabas18a.html>
- [6] Michael Barba and Noah Baustin. 2023. Are San Francisco Police Officers Misreporting the Races of People They Stop? *The San Francisco Standard* (Sept. 2023). <https://sfstandard.com/2023/09/13/san-francisco-police-officer-misrepresented-race-bias-investigation/>
- [7] Ruha Benjamin. 2016. Informed Refusal: Toward a Justice-based Bioethics. *Science, Technology, & Human Values* 41, 6 (Nov. 2016), 967–990. <https://doi.org/10.1177/0162243916656059> Publisher: SAGE Publications Inc.
- [8] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. John Wiley & Sons.
- [9] T. Bowell. [n.d.]. Feminist Standpoint Theory | Internet Encyclopedia of Philosophy. <https://iep.utm.edu/fem-stan/>
- [10] Sarah Brayne. 2020. *Predict and Surveil: Data, Discretion, and the Future of Policing*. Oxford University Press.
- [11] Meredith Broussard, Nicholas Diakopoulos, Andrea L. Guzman, Rediet Abebe, Michel Dupagne, and Ching-Hua Chuan. 2019. Artificial Intelligence and Journalism. *Journalism & Mass Communication Quarterly* 96, 3 (Sept. 2019), 673–695. <https://doi.org/10.1177/1077699019859901> Publisher: SAGE Publications Inc.
- [12] Isabelle Bruno, Emmanuel Didier, and Tommaso Vitale. 2014. Statactivism: Forms of Action between Disclosure and Affirmation. <https://papers.ssrn.com/abstract=2466882>
- [13] Alain Desrosières. 1998. *The politics of large numbers: A history of statistical reasoning*. Harvard University Press.
- [14] Catherine D'Ignazio. 2024. *Counting Feminicide: Data Feminism in Action*. The MIT Press, Cambridge, Massachusetts.
- [15] Catherine D'Ignazio and Lauren F Klein. 2023. *Data feminism*. MIT press.
- [16] Rosalind Edwards and Janet Holland. 2013. *What is Qualitative Interviewing?* Bloomsbury Academic. <https://doi.org/10.5040/9781472545244> Accepted: 2022-10-14T14:53:13Z.
- [17] Danielle Ensign, Sorelle Friedler, Scott Neville, Carlos Scheidegger, and Suresh Venkatasubramanian. 2018. Runaway Feedback Loops in Predictive Policing. *Proceedings of Machine Learning Research* 81 (Jan. 2018). [https://scholarship.haverford.edu/compsci\\_facpubs/113](https://scholarship.haverford.edu/compsci_facpubs/113)
- [18] Wendy Nelson Espeland and Berit Irene Vannebo. 2007. Accountability, quantification, and law. In *Annual Review of Law and Social Science*, John Hagan, Kim Lane Scheppele, and Tom Tyler (Eds.). 21–43. <https://doi.org/10.1146/annurev.lawsocsci.2.081805.105908>
- [19] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Publishing Group.
- [20] Tobias Fahse, Viktoria Huber, and Benjamin van Giffen. 2021. Managing Bias in Machine Learning Projects. In *Innovation Through Information Systems (Lecture Notes in Information Systems and Organisation)*, Frederik Ahlemann, Reinhard Schütte, and Stefan Stieglitz (Eds.). Springer International Publishing, Cham, 94–109. [https://doi.org/10.1007/978-3-030-86797-3\\_7](https://doi.org/10.1007/978-3-030-86797-3_7)
- [21] Sina Fazelpour and David Danks. 2021. Algorithmic bias: Senses, sources, solutions. *Philosophy Compass* 16, 8 (2021), e12760. <https://doi.org/10.1111/phc3.12760> <https://onlinelibrary.wiley.com/doi/pdf/10.1111/phc3.12760>
- [22] Todd Feathers. 2021. Gunshot-Detecting Tech Is Summoning Armed Police to Black Neighborhoods. <https://www.vice.com/en/article/88nd3z/gunshot-detecting-tech-is-summoning-armed-police-to-black-neighborhoods>
- [23] J. Ferguson and Deborah Witzburg. 2021. The Chicago Police Department's Use of Shotspotter Technology. *Chicago, IL: City of Chicago Office of Inspector General* (2021).
- [24] Matt Franchi, J.D. Zamfirescu-Pereira, Wendy Ju, and Emma Pierson. 2023. Detecting disparities in police deployments using dashcam data. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAcCT '23)*. Association for Computing Machinery, New York, NY, USA, 534–544. <https://doi.org/10.1145/3593013.3594020>
- [25] Bruce B. Frey. 2018. *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation*. SAGE Publications, Inc. <https://doi.org/10.4135/9781506326139>
- [26] Timnit Gebru. 2019. Oxford Handbook on AI Ethics Book Chapter on Race and Gender. <https://doi.org/10.48550/arXiv.1908.06165> arXiv:1908.06165 [cs].
- [27] Andre Gorz. 1968. Strategy for labor. A radical proposal. *Science and Society* 32, 4 (1968).
- [28] Ben Green. 2019. Good" isn't good enough. In *Proceedings of the AI for Social Good workshop at NeurIPS*, Vol. 17.
- [29] Ben Green. 2021. Data Science as Political Action: Grounding Data Science in a Politics of Justice. *Journal of Social Computing* 2, 3 (Sept. 2021), 249–265. <https://doi.org/10.23919/JSC.2021.0029> arXiv:1811.03435 [cs].
- [30] Ben Green. 2022. Escaping the Impossibility of Fairness: From Formal to Substantive Algorithmic Fairness. *Philosophy & Technology* 35, 4 (Oct. 2022), 90. <https://doi.org/10.1007/s13347-022-00584-6>
- [31] Ben Green and Salomé Viljoen. 2020. Algorithmic realism: expanding the boundaries of algorithmic thought. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAcT\* '20)*. Association for Computing Machinery, Barcelona, Spain, 19–31. <https://doi.org/10.1145/3351095.3372840>
- [32] Gabriel Grill, Fabian Fischer, and Florian Cech. 2023. Bias as Boundary Object: Unpacking The Politics Of An Austerity Algorithm Using Bias Frameworks. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAcCT '23)*. Association for Computing Machinery, New York, NY, USA, 1838–1849. <https://doi.org/10.1145/3593013.3594120>
- [33] Bernard E. Harcourt. 2008. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. University of Chicago Press.
- [34] Sandra Harding. 1992. Rethinking Standpoint Epistemology: What is "Strong Objectivity?". *The Centennial Review* 36, 3 (1992), 437–470. <https://www.jstor.org/stable/23739232>
- [35] Kashmir Hill. 2023. *Your Face Belongs to Us: The Secretive Startup Dismantling Your Privacy*. Simon and Schuster.
- [36] Anna Lauren Hoffmann. 2019. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society* 22, 7 (2019), 900–915.
- [37] Bell Hooks. 2000. *Feminist theory: From margin to center*. Pluto Press.
- [38] Alec Karakatsanis. 2019. *Usual cruelty: The complicity of lawyers in the criminal injustice system*. The New Press.
- [39] Kari Karppinen and Hallvard Moe. 2011. What we talk about when we talk about document analysis. In *Trends in Communication Policy Research*. Intellect, Bristol.
- [40] Leslie W. Kennedy, Joel M. Caplan, and Eric L. Piza. 2018. *Risk-Based Policing: Evidence-Based Crime Prevention with Big Data and Spatial Analytics*. Univ of California Press.
- [41] Yana Kunichoff and Patrick Sier. 2017. The Contradictions of Chicago Police's Secretive List. <https://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/>
- [42] Michael Lipton. 1977. *Why poor people stay poor : a study of urban bias in world development*. Temple Smith ; Australian National University Press. <https://openresearch-repository.anu.edu.au/handle/1885/114902> Accepted: 2017-04-18T05:59:25Z Last Modified: 2022-11-17.
- [43] Kristian Lum and William Isaac. 2016. To predict and serve? *Significance* 13, 5 (2016), 14–19. <https://doi.org/10.1111/j.1740-9713.2016.00960.x> <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x>
- [44] Steve Mann, Jason Nolan, and Barry Wellman. 2003. Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments. *Surveillance & society* 1, 3 (2003), 331–355.
- [45] Sandra G. Mayson. [n.d.]. Bias In, Bias Out. <https://www.yalelawjournal.org/article/bias-in-bias-out>
- [46] Allegra M McLeod. 2018. Envisioning abolition democracy. *Harv. L. Rev.* 132 (2018), 1613.
- [47] William B. Meyer and Jessica K. Graybill. 2016. The suburban bias of American society? *Urban Geography* 37, 6 (Aug. 2016), 863–882. <https://doi.org/10.1080/02723638.2015.1118990> Publisher: Routledge <https://doi.org/10.1080/02723638.2015.1118990>
- [48] Matthew B. Miles, A. Michael Huberman, and Johnny Saldana. 2014. *Qualitative Data Analysis*. SAGE.
- [49] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. Model Cards for Model Reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAcT\* '19)*. Association for Computing Machinery, New York, NY, USA, 220–229. <https://doi.org/10.1145/3287560.3287596>
- [50] Diba Mohtasham. 2024. Chicago will drop controversial ShotSpotter gunfire detection system. *NPR* (Feb. 2024). <https://www.npr.org/2024/02/15/1231394334/shotspotter-gunfire-detection-chicago-mayor-dropping>
- [51] Bryce Newell. 2020. Introduction: The State of Sousveillance. *Surveillance & Society* 18, 2 (June 2020), 257–261. <https://doi.org/10.24908/ss.v18i2.14013>
- [52] Amelia Nierenberg. 2023. Over 100 Connecticut State Troopers Accused of Faking Traffic Stops. *The New York Times* (Sept. 2023). <https://www.nytimes.com/2023/09/04/nyregion/connecticut-false-tickets.html>
- [53] C. Parker, S. Scott, and A. Geddes. 2019. Snowball Sampling. *SAGE Research Methods Foundations* (Sept. 2019). <http://methods.sagepub.com/foundations/snowball-sampling> Publisher: SAGE.
- [54] Samir Passi and Solon Barocas. 2019. Problem formulation and fairness. In *Proceedings of the conference on fairness, accountability, and transparency*. 39–48.
- [55] Beth Pearsall. 2010. Predictive policing: The future of law enforcement. *National Institute of Justice Journal* 266, 1 (2010), 16–19.
- [56] Theodore M. Porter. 1995. *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton University Press, Princeton, NJ.
- [57] Dasha Pruss. 2023. Ghosting the Machine: Judicial Resistance to a Recidivism Risk Assessment Instrument. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*. 312–323.
- [58] Paola Ricaurte. 2019. Data epistemologies, the coloniality of power, and resistance. *Television & New Media* 20, 4 (2019), 350–365.
- [59] David Richards. 1996. Elite Interviewing: Approaches and Pitfalls. *Politics* 16, 3 (1996), 199–204. <https://doi.org/10.1111/j.1467-9256.1996.tb00039.x> <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9256.1996.tb00039.x>

- [60] Rashida Richardson, Jason Schultz, and Kate Crawford. 2019. *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*. SSRN Scholarly Paper ID 3333423. Social Science Research Network, Rochester, NY. <https://papers.ssrn.com/abstract=3333423>
- [61] Victor Rotaru, Yi Huang, Timmy Li, James Evans, and Ishanu Chattopadhyay. 2022. Event-level prediction of urban crime reveals a signature of enforcement bias in US cities. *Nature Human Behaviour* 6, 8 (Aug. 2022), 1056–1068. <https://doi.org/10.1038/s41562-022-01372-0> Number: 8 Publisher: Nature Publishing Group.
- [62] Kathryn Roulston and Myungwon Choi. 2018. *The SAGE Handbook of Qualitative Data Collection*. SAGE Publications Ltd. <https://doi.org/10.4135/9781526416070>
- [63] Andrew D Selbst. 2017. Disparate impact in big data policing. *Ga. L. Rev.* 52 (2017), 109.
- [64] Aaron Shapiro. 2019. Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing. *Surveillance & Society* 17, 3/4 (Sept. 2019), 456–472. <https://doi.org/10.24908/ss.v17i3/4.10410>
- [65] Maya Simkin. 2023. Looking Forward to the End of Chicago’s Contract with ShotSpotter. <https://www.chicagoappleseed.org/2023/05/10/ending-chicago-contract-with-shotspotter/>
- [66] Mario Luis Small. 2009. ‘How many cases do I need?’: On science and the logic of case selection in field-based research. *Ethnography* 10, 1 (March 2009), 5–38. <https://doi.org/10.1177/1466138108099586> Publisher: SAGE Publications.
- [67] Katherine B Spencer, Amanda K Charbonneau, and Jack Glaser. 2016. Implicit bias and policing. *Social and Personality Psychology Compass* 10, 1 (2016), 50–63.
- [68] Megan T. Stevenson. 2018. Assessing Risk Assessment in Action. *Minnesota Law Review* 103 (2018), 303. <https://papers.ssrn.com/abstract=3016088>
- [69] Matt Stroud. 2021. An automated policing program got this man shot twice. <https://www.theverge.com/c/22444020/heat-listed-csk-entry>
- [70] Annie Sweeney. 2020. For years Chicago police rated the risk of tens of thousands being caught up in violence. That controversial effort has quietly been ended. <https://www.chicagotribune.com/2020/01/24/for-years-chicago-police-rated-the-risk-of-tens-of-thousands-being-caught-up-in-violence-that-controversial-effort-has-quietly-been-ended/>
- [71] Eric Umansky. 2023. The Failed Promise of Police Body Cameras. *The New York Times* (Dec. 2023). <https://www.nytimes.com/2023/12/13/magazine/police-body-cameras-miguel-richards.html>
- [72] Maarten Van Craen and Wesley G Skogan. 2017. Achieving fairness in policing: The link between internal and external procedural justice. *Police quarterly* 20, 1 (2017), 3–23.
- [73] Pransha Verma. 2023. Heat Listed. *The Washington Post* (Dec. 2023). <https://www.washingtonpost.com/technology/2022/07/15/predictive-policing-algorithms-fail/>
- [74] Richard A Webster. 2021. “If everybody’s white, there can’t be any racial bias”: The disappearance of hispanic drivers from traffic records. *ProPublica—Journalism in the Public Interest*. <https://www.propublica.org/article/if-everybodys-white-there-cant-be-any-racial-bias-the-disappearance-of-hispanic-drivers-from-traffic-records> (2021).
- [75] Robert Stuart Weiss. 1994. *Learning from Strangers: The Art and Method of Qualitative Interview Studies*. Free Press. Google-Books-ID: SLM6AAAAIAAJ.
- [76] Michael D White, Henry F Fradella, and Michaela Flippin. 2021. How can we achieve accountability in policing? The (not-so-secret) ingredients to effective police reform. *Lewis & Clark L. Rev.* 25 (2021), 405.
- [77] David Gray Widder and Dawn Nafus. 2023. Dislocated accountabilities in the “AI supply chain”: Modularity and developers’ notions of responsibility. *Big Data & Society* 10, 1 (2023), 20539517231177620.
- [78] David Gray Widder, Dawn Nafus, Laura Dabbish, and James Herbsleb. 2022. Limits and possibilities for “Ethical AI” in open source: A study of deepfakes. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. 2035–2046.
- [79] Douglas Yeung, Inez Khan, Nidhi Kalra, and Osonde A. Osoba. 2021. *Identifying Systemic Bias in the Acquisition of Machine Learning Decision Aids for Law Enforcement Applications*. Technical Report. RAND Corporation. <https://www.rand.org/pubs/perspectives/PEA862-1.html>
- [80] Brian Hu Zhang, Blake Lemoine, and Margaret Mitchell. 2018. Mitigating Unwanted Biases with Adversarial Learning. <https://doi.org/10.48550/arXiv.1801.07593> arXiv:1801.07593 [cs].
- [81] Miri Zilka, Riccardo Fogliato, Jiri Hron, Bradley Butcher, Carolyn Ashurst, and Adrian Weller. 2023. The Progression of Disparities within the Criminal Justice System: Differential Enforcement and Risk Assessment Instruments. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAcCT ’23)*. Association for Computing Machinery, New York, NY, USA, 1553–1569. <https://doi.org/10.1145/3593013.3594099>

## A INTERVIEW GUIDE

### (1) Context and Positionality

- Can you introduce yourself?
- Can you tell us more about your work?
- How are you related to the predictive policing project? How have you heard of it, if at all?
- How are you involved with other initiatives around community and public safety in Chicago?

### (2) Bias definition and utility

- What is algorithmic bias to you in general?
- What is the problem with it? Why do you think it exists?
- What is enforcement bias to you specifically?
- What is the problem with it? Why do you think it exists?
- What is its specific relevance to your work? And the context of Chicago?
- What is your role in tackling bias and problems faced in doing so (if applicable) at work?

### (3) Perceived feasibility and desirability of using evidence of bias to tackle social disparities

- How should bias be addressed?
- What would an ideal solution look like to you (especially with relevance to your role/work)?
- What are the barriers to resolving bias? How can it be realistically tackled, in your opinion?

### (4) Snowball sampling

- Do you know other people it might be relevant for us to interview?

## B CODE TABLE

### Abolition

---

Alternatives to current system/police  
 Building healthy communities  
 Calls for changing the narrative/status quo  
 The need to acknowledge harms of current system  
 Importance of history and education  
 Importance of hope and belief in change

### Co-option and exploitation

---

Academic exploitation  
 Co-opting evidence-based label  
 Commercial interest in predicting crime  
 Disruption rhetoric  
 Open source as a marketing strategy  
 Responsibilities of developers

### Downsides of AI

---

AI is racist/not objective  
 Inherent reservations about AI  
 Garbage in garbage out/distrust of data  
 Human discretion is important  
 Nuance not captured by algorithm  
 Technology isn’t going to fix policing

### Lived experience and positionality

---

Affected people are best suited to solve problems  
 Barriers to entering research  
 Impacted groups lack resources  
 Different stakeholders care about different forms of bias  
 People not affected don’t have the same understanding  
 Who generated the data/designed the tool matters

Importance of trust

---

**Reformism**

Interpersonal bias rather than structural

Resource allocation

Safety is the priority

Police promotes safety

---

**Structural bias**

Importance of addressing root cause of crime

Racial profiling and segregation

Structural racism requires a structural solution

The system is biased

---

**Upsides of AI**

Algorithm can be more objective/humans are biased

Algorithms increase efficiency

Open source

Using data to understand authority figures (sousveillance)

Information has multiple uses

---