Title: 5 Challenges Facing Data-Driven Policing

Authors:
Duncan Purves, Assistant Professor of Philosophy, University of Florida (Corresponding author)
Contribution: Conceptualization; funding acquisition; investigation; methodology; project administration; resources; writing-original draft; writing-review and editing
Jeremy Davis, Postdoctoral Associate in Philosophy, University of Florida
Contribution: Conceptualization; investigation; supervision; methodology; project administration; resources; writing-original draft; writing-review and editing
Juan Gilbert, The Banks Family Preeminence Endowed Professor of Computer and Information Science and Engineering, University of Florida
Contribution: Conceptualization; writing-review and editing; supervision

Abstract: The use of data-driven approaches by law enforcement is at an inflection point. Heralded in the 1990s as the next great revolution in police work, mounting criticism has led police departments to rein in their ambitions of fighting crime using big data and computer algorithms. This paper outlines several data-driven policing initiatives currently in use in the United States, before laying out five challenges facing data-driven policing: Racial bias and civil liberties; what data for which purposes; standards of success; explanation and opacity; accountability and community oversight. We describe several recent attempts to address these challenges, highlighting their advantages and shortcomings. Ultimately, responsible data-driven policing requires collaboration between communities, academics, technology developers, police departments, and policy-makers to confront and address these challenges. And as we will see, it may also require careful democratic deliberation about the role and value of law enforcement in society.

Key words: data-driven policing; fairness; transparency; accountability; bias; ethics

Full manuscript content:

1 Introduction:

The use of data-driven approaches by law enforcement is at an inflection point. Heralded in the 1990s as the next great revolution in police work, mounting criticism has led police departments to rein in their ambitions of fighting crime using big data and computer algorithms (Heaven 2020). After years of local resistance and national criticism about potential racial bias, the Los Angeles Police Department, in the spring of 2020, ended its decade-long experiment with PredPol, an algorithmic crime forecasting system designed to predict the location of property crimes (Haskins 2019; Moravec 2019). Recently, many others have followed suit: in June of 2020, Santa Cruz, California became the first U.S. city to ban predictive policing technology (Sturgill 2020). In January of 2021, the city council in Oakland, California voted unanimously to ban both predictive policing and biometric surveillance technology (Burbank 2021). Similar bans in other cities across the country are sure to follow.

In the meantime, data-driven policing methods continue to be used by dozens of police departments across the country, and their use appears to be growing at a fast pace. Given the constantly evolving landscape for data-driven policing methods and the inconsistent publicity about the use of such technologies, it is difficult to know precisely how many departments employ data-driven methods, but it is safe to say that most major cities across the U.S., as well as several smaller cities and towns, use one or more forms of data-driven policing technology. Furthermore, even some departments that have discontinued using certain methods still rely on others. For example, despite abandoning PredPol, the LAPD Strategic Plan for 2019-2021 lists "enhancing data-driven policing" as the #1 activity in its initiative to reduce crime and victimization (Los Angeles Police Department 2019).

In what follows, we begin by outlining several data-driven policing initiatives currently in use in the United States, before laying out five challenges facing data-driven policing. Some challenges have received considerable attention already. Others are mostly overlooked. Ultimately, responsible data-driven policing requires collaboration between communities, academics, technology developers, police departments, and policy-makers to confront and address these challenges. And as we will see, it may also require democratic deliberation about the role and value of police in society.

2 What is data-driven policing?

Data-driven policing is part of a larger "big data" revolution across sectors, including medicine, finance, transportation, and criminal justice. What characterizes "big data", and hence, data-driven policing, is that it takes vast quantities of data from a variety of data sources, and, using rapid computer processing and algorithms, identifies correlations or patterns in the data that humans would otherwise overlook, in order to classify new instances in the domain of interest (Brayne 2020; Ferguson 2017, pp. 8–9; Lazer and Radford 2017). In many instances, this process of pattern recognition and classification occurs through the use of computer algorithms, which are sometimes developed through the process of machine learning. Data-driven policing refers to a variety of big data applications in law enforcement. It is best characterized by way of examples.

The LAPD pioneered perhaps the most widely critiqued data-driven policing application. Starting in 2011 the LAPD began an experiment with PredPol. PredPol was developed by anthropologist Jeffrey Brantingham and computer scientist George Mohler. It predicts crime using a proprietary computer algorithm trained on data about the type, location, and time of crime. Premised on a criminological phenomenon known as the "near repeat effect," wherein one crime event in an area tends to give rise to a spike in similar crimes in that area, PredPol was used by the LAPD to forecast future property crime—vehicular theft and theft from a vehicle—at highly specific locations and times during a police officer's patrol shift (Short et al. 2009). These forecasts are displayed on a computer screen as "heat maps" of 500x500 square foot boxes indicating high-risk areas in the officer's beat. Additional information about time, location, and type of recent crime incidents can be incorporated each day to update the algorithm's forecasts. While the LAPD discontinued its use of PredPol in the spring of 2020, PredPol has become one of the most widely used pieces of predictive policing software in the United States (Miller 2020).

Whereas PredPol was primarily used by the LAPD to assist in-house decision-making about where police officers on patrol spent their uncommitted time, other data-driven systems are more holistic in their application. Risk Terrain Modeling (RTM) is a crime forecasting system developed by Rutgers criminologists Joel Caplan and Leslie Kennedy. RTM is theoretically grounded in so-called "opportunity theory," which explains criminal activity in terms of the social and physical features of places that give rise to opportunities to commit crime (Cohen et al. 1981). To understand where opportunities for crime are most abundant, one must be able to identify the features of places that give rise to those opportunities. As such, RTM forecasts crime at a location by incorporating in its model the features of a location that are the underlying causes of crime. Once the features of locations that give rise to crime are identified, geographic units are assigned a value depending on the presence or absence of those features. The higher the value at a location the greater the risk of crime at that location. The key difference between RTM and PredPol is that RTM forecasts take into account both crime data and information about the features of a place that make it vulnerable to crime, whereas PredPol merely makes use of past crime data. In one project, RTM was used to predict the location of shootings in New Jersey. Spatial features used to predict shootings included the addresses of known gang members, retail businesses and drug arrests for sale or possession (Caplan et al. 2011).

Still other data-driven policing methods are designed not to predict high-risk places but high-risk people. Around the same time that the LAPD began its experiment with PredPol, it adopted another predictive policing method called "Operation LASER," short for Los Angeles Strategic Extraction and Restoration Program (Uchida et al. 2012). Part of Operation LASER involved identifying a list of "chronic offenders." The list of chronic offenders was compiled using information gathered from patrol units, field interview cards, traffic citations, crime and arrest reports and criminal histories. Chronic offenders are ranked by a points system. Points-accruing features included having a violent crime conviction, a known gang affiliation, having been arrested for possession of a handgun, and police contacts. From this list, the LAPD generated "Chronic Offender Bulletins," which included information about the individual's arrest history, physical markers, gang affiliation, parole status, warrants, and recent police contacts, among other information. These bulletins were circulated within departments so as to inform officers on patrol about which individuals required the most attention (Brayne 2020). After an internal audit in 2019 revealed gross inconsistencies in the application of Operation LASER by officers in the field, LAPD discontinued the program.

However, similar person-based policing programs remain in use by police departments across the country. For example, the Tampa Bay Times recently revealed some of the inner workings of Pasco County Sheriff's Office's Intelligence Led Policing program (McGrory and Bedi 2020). Part of the Sheriff's Office's intelligence-led policing approach is to look for "prolific offenders" in Pasco County, described in the intelligence-led policing manual as individuals who are "not likely to reform." The prolific offender designation is determined in part by using an algorithm to assign points to members of the community if they have been arrested for a violent crime, narcotics violation, theft, or if they are *suspected* of any such offenses. Point "enhancements" are assigned for involvement in five or more crime reports either as a reporting person, victim, or

witness. The more points a person has, the higher risk they are deemed to be of committing a future crime. District analysts review the points assigned to individuals by the algorithm and may designate them as prolific offenders. These individuals are then singled out for greater police scrutiny. We say more about Pasco County's Intelligence Led Policing program below.

Having laid out some common data-driven policing methods and initiatives, we now turn to several of the most significant challenges that the use of such methods poses. Several of these challenges arise from the implementation of the methods, while other challenges call into question some of the goals of policing that the methods presuppose.

3 Challenge #1 Racial bias and civil liberties

The most prominent objection to data-driven policing found in academic and media publications is that it leads to discriminatory treatment against people of color, or against economically disadvantaged classes. This discriminatory treatment can result from a combination of two flaws in some predictive policing systems. First, the data used to predict high-risk places and people can have dubious origins. Differential selection theory suggests that arrest data, particularly for drug and nuisance crimes, is significantly influenced by racial bias in police officers' choices about whom to investigate (Piquero and Brame 2008). If arrests are a reflection of racial bias on the part of police, and arrests are used to generate the forecasts of predictive policing systems, then these systems will tend to forecast greater crime in minority communities than actually exists. Thus, residents of those communities will often face unnecessary, and sometimes unwanted, additional police attention. Second, once police inundate an area, this can lead to even more police contacts, incidents, and arrests in minority communities. This data is fed back into the system and used to make future predictions, leading to a "ratchet effect" of escalating police attention (Harcourt 2006). Insofar as the predictive success of predictive policing systems hinges on ratcheting up the unjust and discriminatory policing patterns of the past, these systems are not ethically justifiable. Nor are they particularly useful.

Whether the concern about discriminatory treatment applies to particular predictive policing systems will depend on a number of factors, including the kind of data the system uses to generate predictions. For example, the LAPD adopted PredPol in part because the system does *not* rely on arrest data to generate crime forecasts. In this way it seems to avoid influence from police officer selection bias. Sean Malinowski of the LAPD describes this rationale in Andrew Guthrie Ferguson's seminal book on constitutional issues related to data-driven policing.

> *...arrests are not part of the equation [in generating PredPols forecasts]. We felt this was important because we heard from some community members that they were concerned about the program creating a kind of self-fulfilling prophecy from under which a community could not recover...In our model, we would hope to deploy the officer based on crime only and then deny the criminal the opportunity to commit the crime in the first place* (Ferguson 2017, p. 74).

Even without arrest data, a ratchet effect can occur, depending on what non-arrest data is used in a predictive model. For example, police reports can come from community members

reporting crimes and from officers looking for crimes, so police reports are susceptible to influence from selection bias. Emergency calls for service, on the other hand, are entirely community-driven, so they offer an attractive alternative to arrest data and crime reports.

Other police departments using data-driven policing systems have seemingly ignored the threat of a ratchet effect altogether. Pasco County's Intelligence Led Policing identifies "prolific offenders" using a variety of data sources, each of which is susceptible to influence by police bias. By assigning points to individuals based on arrests and mere *suspicion* of criminality, the Pasco County system faces a clear threat from differential selection. By "enhancing" point assignments for mere involvement in crime reports, either as a reporting person, victim, or witness, the system risks placing undue burden on innocent members of the community.

Perhaps the most remarkable fact about the public and academic conversation around data-driven is how little evidence exists either of bias and discriminatory impacts, or of data-driven policing's efficacy in preventing crime. Only a small handful of studies have been published assessing the efficacy of data-driven programs (Boba Santos 2020), and only *one* published study has examined the racially discriminatory impact of PredPol on arrest rates (Brantingham and Valasik 2018), finding no statistically significant effect on racial disparities in arrests. Of course, absence of evidence is not evidence of absence, particularly when access to data-driven policing systems is denied to the public or to academic investigators, as it often is (Boba Santos 2020).

Something else that has been lost in the discussions of bias and discriminatory impacts is how to better include members of minority communities in decision-making about data-driven systems and additional police attention being directed toward their community. No studies have been published about public attitudes toward the use of algorithms to predict where crime will occur or who will commit it. And the evidence from national surveys paints a nuanced picture of minority attitudes toward police presence in their communities. Black Americans express much less favorable attitudes toward police than do White Americans (Ekins 2016), and black Americans have a long-documented distrust of the legal system and of police in particular (Bobo and Thompson 2006; Brooks 2000; Hagan and Albonetti 1982). And yet in a recent Gallup poll, 61% of Black Americans want police presence in their area to remain the same, and 19% wanted police to spend *more* time in their area (Saad 2020). The details of when, how, and what sort of police attention in minority communities is an unwanted *burden* is a complicated matter.

Person-based predictive policing programs like Pasco County's prolific offender program raise a host of distinct challenges relating to civil liberties and surveillance. What academics and journalists have found most troubling about these programs is that they often target individuals who are not suspected of any particular crime, but, because of their social networks, have become *associated* with criminal activity. But once police have fixed their gaze on a so-called prolific offender, and information about their prolific offender status circulates to other institutions, it can become virtually impossible to get back onto the right side of the law (Brayne 2020).

When data about criminal justice contacts is integrated and shared across institutions, as it increasingly is, one consequence of labeling some individuals as chronic or prolific offenders is that people with records of contact with the criminal justice system may shy away from other institutions that keep formal records. Sociologist Sarah Brayne has argued that individuals with police contacts,

arrests, or convictions are less likely to participate in medical, financial, or educational institutions, or the labor market than individuals who do not have those recorded contacts with the criminal justice system (Brayne 2014). Thus, what begins as an inequality in the criminal justice system can entrench inequality in access to other vital social services.

Being marked by an algorithm as a prolific offender has implications for constitutional protections as well. The fourth amendment protects citizens from unreasonable stops, searches, and seizures by law enforcement. Whether a stop counts as reasonable depends in part on whether a police officer had reasonable suspicion that the subject of the search was involved in a criminal activity. As Andrew Ferguson has noted in several places, what constitutes reasonable suspicion for a search might be affected by algorithmic classifications (Ferguson 2015). Behavior that would otherwise be insufficiently suspicious to justify a search may warrant a search if the officer observing the person knows that he or she has been flagged as a "prolific offender" by a predictive policing program. But this additional suspicion has nothing to do with the *current activities* of the suspect. The suspicion is grounded in a statistical inference made by a computer algorithm.

Addressing discrimination and violations of civil liberties in data-driven policing requires a multi-pronged approach, and it cannot be achieved over night. As we've suggested, care must be taken in data source selection and processing to avoid a ratchet effect. Beyond that, continuous monitoring of the impacts of data-driven systems on communities is required to guard against disparate impacts. This has not been done effectively in the past. One of the key deficiencies of the PredPol program identified in the Office of the Inspector General's audit of LAPD's data-driven policing programs was inconsistent recording by officers of time spent in the designated high-risk zones (Los Angeles Police Commission 2019). The program had been in effect for approaching ten years by the time the OIG made this discovery.

4 Challenge #2 What data for which purposes?

Our second challenge concerns two assumptions made by proponents of data-driven policing that its critics have forcefully called into question.

4.1 Assumption 1 advocates of data-driven policing often tacitly assume that once a department adopts certain data-driven methods, these methods can be used in a value-neutral way. That is, it is taken as obvious how a given technology is to be incorporated into existing practices, what roles or tasks it will assume, and what the parameters of its function will be. But a growing body of academic literature is calling this assumption into question (Barocas and Selbst 2016; Brayne 2020; Selbst 2017).

We just saw, in the previous section, that data-driven policing methods have the potential to further entrench various discriminatory elements of the criminal justice system. But problems like these also result from decisions about where in a given department's strategic decision-making process these technologies are introduced.

To begin, take the case of place-based predictive policing. There are several junctures in the decision-making process where this technology can be employed. First, it could be consulted at an early stage, to determine where to distribute the department's resources at a given time—i.e., which

macro-level places (e.g., beats, city wards, districts, or neighborhoods) ought to be prioritized. Alternatively, the system could be consulted after having made data-independent judgments about how to allocate police resources among macro-level places. At this point, the predictive software could help focus the officers' attention on a given block or address within the broader area they have selected.

The choice between these two options may seem governed entirely by strategic considerations; however, the evaluative assumptions embedded in this decision raise important ethical issues. In particular, while the former option—consulting the data at the outset—seems more efficient, it is also more likely to result in targeting those in minority communities that have been historically and systemically burdened by excessive police scrutiny. This, after all, is the clearest example of how data-driven policing reproduces the history of racist policing. The second option—only consulting the data at later stages—precludes this: entire communities would not be subject to police scrutiny on the basis of an algorithm. But even this alternative could potentially engender problematic outcomes—for example, if officers were biased in their initial selection of areas to patrol.

Different technological systems also suggest different solutions to crime. For example, PredPol uses data only on timing, location, and type of crime. Forecasts based on these data can predict when and where crime will occur, but they cannot diagnose the underlying causes of crime. For this reason, such a system lends itself to a patrol- or enforcement-oriented response to crime. If all a police department knows is when and where the crime is likely to occur, the natural response is to send patrol officers to the location in order to deter or apprehend the offender. Compare this system with one like RTM that incorporates data from non-law enforcement agencies about features of high crime places. Such a system might, for example, find correlations between poor street lighting or multi-family housing and auto vehicle theft. But here the system has moved away from crime prediction to diagnosis, and it therefore suggests different non-enforcement-oriented solutions. Addressing the underlying features of places that make them vulnerable to crime might require engaging non-law-enforcement agencies like public works, sanitation, or urban planning. Therefore, choices about which data to use in data-driven policing implicitly or explicitly involve choices about the proper role of police in crime prevention.

Further, much of the focus of big data policing systems centers on 'street crime'—e.g., property damage, theft, assault, and so forth. This focus is a choice on the part of police departments and those responsible for designing and programming these systems. This can be seen most clearly in the relative lack of emphasis on other forms of crime—for example, what is often referred to as 'white-collar crime', as well as other types of harmful and socially disruptive criminal activity, and yet is not often actively pursued by local police departments. There are several possible explanations for this focus, such as the significant data available on street crime, which better enables the discovery of patterns within the data that facilitates future predictions. Or it could be attributed to political pressures, departmental expertise, and so forth. In any case, it is important to recognize that, as a result of this narrower focus, the measurements of success or failure of data-driven systems will be sensitive only to their ability to thwart this particular subset of crimes. This leads to several underappreciated issues, such as the creation of incentives for police departments to

increase focus on crimes disproportionately committed by the poor; the increased criminalization of certain minor offenses; and the heightened potential for harassment of those with prior criminal activity, and unwanted attention to those in certain geographic areas (Barocas and Selbst 2016).

These examples illustrate several aspects of the broader ethical concern relating to implementation of these data-driven methods. More generally, police departments implementing such technologies face the following challenge: What data should we use, and for what purposes should we use it? Much of the discussion surrounding the implementation of data-driven policing systems has thus far ignored this question. As a first step, departments planning to adopt such methods ought to scrutinize some of the values and priorities that adoption assumes. And this scrutiny must not occur in a vacuum. Insofar as data-driven policing systems encourage more aggressive enforcement-oriented policing tactics than other systems, community members have a significant stake in participating in the deliberative process about which systems are adopted. For example, citizen-led police advisory councils must be empowered to communicate to police the needs and concerns of community members. As we discuss in Challenge #4 below, this will require greater transparency about the role that data-driven systems play in policing.

4.2 Assumption 2 The second major assumption that is taken for granted by advocates of data-driven policing, and which we have been granting so far, is that the primary function of data-driven tools in policing is to promote social well-being. A growing chorus of critics has called this assumption into question by calling into question the social value of police (Adams and Rameau 2016; Vitale 2017). Calls to abolish or defund the police have grown in volume and frequency over the past decade as killings of unarmed Black men and women by police continue to make national headlines. If policing is not a legitimate institution, then the tools it uses are not legitimate either. Police abolitionists have argued that the institution of policing is, and always has been, an unjust mechanism of social control of minority bodies, and that one need only glimpse the history of policing to see that this is true (Muhammad 2019). On this view, data-driven policing tools only serve to maintain that control (Capers 2017). Indeed, some of the staunchest opponents of data-driven policing are explicitly abolitionist. Hamid Khan is the director of the Stop LAPD Spying Coalition, the leading anti-data-driven activist group in Los Angeles. In interviews he has described the group as "fiercely abolitionist" (Ryan-Mosley and Strong 2020). Khan's opposition to data-driven technologies like PredPol and Operation Laser is founded in a commitment to the idea that the system of which these tools are a part is irredeemably racist.

Opposition to data-driven policing that is rooted in the police abolitionist movement leaves little room for reform. From this perspective, any tool of the police will, necessarily, be used for oppressive purposes and so must be dismantled. Data-driven technologies are a tool of the police. Therefore they must be dismantled. Whether or not one accepts the premise of the abolitionist movement, the critique raises a vital complication for the question we started with: "which data for which purposes?" If existing data-driven policing technologies are tools of oppression and racial control, then we shouldn't be using data *at all* in the service of these technologies (Noble 2018). Instead what data we use, and how we use it, requires reimagining the fundamental goals guiding technology development. Ruha Benjamin is a leading academic voice of this perspective, arguing

that the social aim of technology should be to promote equity and social justice, and not be driven by concerns for maximal efficiency and profit motive (Benjamin 2019).

5 Challenge #3 Standards of success

As we noted above, a common refrain among critics of predictive policing is that predictive policing systems will discriminate against people of color. However, human decision-makers are far from perfect when it comes to being influenced by racial bias. This raises a key question: when evaluating a data-driven system's accuracy, transparency, or fairness, what is the relevant standard of comparison?

This question faces applications of algorithms and machine learning across all sectors, but it arises in high relief in criminal justice contexts where the stakes are very high. Is it good enough for data-driven policing systems to be more accurate, transparent, and fair than the status quo in policing, or must they be held to a higher standard? While the algorithms will never be perfect, neither is the status quo. And waiting too long to adopt the technology might allow inferior policing operations to prevail in the meantime.

To illustrate the issue at stake, let us consider Chicago's experiment with a person-based predictive policing system called "Strategic Subject's List," also known as "The Heat List." The Heat List, now defunct, used an algorithm to identify individuals at high risk of perpetrating or being victimized by a crime. The algorithm assigned rankings to individuals on the basis of certain features, including arrests for drugs, illegal possession of a firearm, assaults, and gang membership. Investigations of this program revealed significant racial disparities in risk assignments. Black men were vastly overrepresented among subjects on the Heat List (Dumke and Main 2017). Does this disparity alone mean that Chicago Police should have stopped using the Heat List to assess risk of criminality and victimization?

Answering this requires first establishing a standard of success for algorithmic tools in criminal justice. One sensible standard is that an algorithmic tool is acceptable to use if it surpasses human decision-making in the same domain of application (Vestby and Vestby 2019). If one aim of these tools is to produce less discriminatory outcomes than human decision-makers, this aim might still be achieved even if The Heat List over represents members of some racial groups. First, some academics specializing in the evaluation of risk assessment tools in criminal courts have argued that tradeoffs between accuracy and fairness apply regardless of whether or not one uses algorithms in the decision-making task (Skeem and Lowenkamp 2020). Other authors have argued that even imperfect algorithmic risk assessment tools are preferable to human decision-makers in the same domain. For only algorithmic systems provide opportunities to identify the *sources* of discriminatory impacts and suggest remedies for those sources. Human decision-makers, whose procedures for arriving at risk-assessment judgments are entirely inaccessible, offer no such remedies. As Jon Kleinberg et al put it, when it comes to addressing discriminatory outcomes, "The opacity of the algorithm does not prevent us from scrutinizing its construction or experimenting with its behavior – two activities that are impossible with humans" (Kleinberg et al. 2018, p. 114). Moreover, they argue, a properly calibrated algorithm can counteract discriminatory intent by human decision-

makers in criminal justice by providing a "counterweight" to their judgment (Kleinberg et al. 2018, pp. 114–115).

One concern with setting human performance as the standard of success is that it might set the bar too low. For example, in the one academic study measuring PredPol's predictive accuracy in the field in comparison with human crime analysts, PredPol's predictions were compared with human crime analysts in Kent, UK and Los Angeles, California. In Los Angeles, PredPol predicted 4.7% of crimes compared with 2.2% predicted by human crime analysts. On one reading of the results, PredPol was a vast improvement over human crime analysts, more than doubling their accuracy. But this feat is only impressive if we think that human crime analysts are already pretty good at predicting where crime is going to occur. Does predicting 2.2% of the locations where crime will occur amount to expertise? And is it worth investing in an algorithmic system that predicts 2.5% more crime than human analysts? As Andrew Ferguson says, describing the study, "To say X is better than Y is only really meaningful if you have a baseline understanding of the value of Y. Maybe both the algorithm and the analyst are terrible, so being better than terrible is not necessarily worth the investment" (Ferguson 2017, p. 70). And as Ferguson notes, there are no academic studies measuring the accuracy of human crime analysts, so it is at present impossible to assess their adequacy as a baseline of comparison.

Brantingham and Valasik's study examining discriminatory arrest patterns when using PredPol found no significant differences in the proportion of arrests by racial-ethnic groups when comparing police patrols allocated by PredPol to patrols allocated by human crime analysts (Brantingham and Valasik 2018). But once again, achieving parity with human decision-makers may set too low a bar for success if arrests themselves are racially biased. If too many Blacks and Latinos are arrested when patrol officers are distributed in accordance with the predictions of human analysts, *maintaining* that proportion of arrests of Blacks and Latinos is not something to be celebrated. We should instead demand that the algorithmic system *improve* on the status quo. And indeed, the differential selection of arrestees by race and ethnicity is a well-established phenomenon for certain types of crime (Beck and Blumstein 2018; Blumstein 2011; Brayne 2020).

From the perspective of the abolitionist, the fact that data-driven policing systems maintain, but fail to improve on, discriminatory patterns of policing is *precisely* why these systems are unacceptable. Because an abolitionist rejects the very aims of the criminal justice institutions in which these data-driven systems are embedded, whatever standard of success we set for these systems will be unsatisfactory.

6 Challenge #4 Explanation and transparency

Like many other algorithmic systems, data-driven policing systems have been criticized for their relative lack of transparency—that is, their opacity. A system is opaque in this way when the contribution of any single feature of the world to the final prediction cannot be easily understood, either by the human decision-maker or the person directly affected by the prediction (Burrell 2016).

Data-driven policing systems are opaque in several different but overlapping ways. First, the algorithms at the core of such systems are extraordinarily technically complex. Certainly, most of those subjected to them are incapable of understanding them. But this is also true of most of the

officers who employ them. In her research on predictive policing, Sarah Brayne spent time alongside LAPD officers as they used this technology in their work. Most of them, she says, admitted that they do not understand how or why the algorithm works.

Furthermore, many data-driven systems are programmed to 'learn' over time—i.e, correct from previous mistakes, and make more accurate assessments. As a result, with increased use, the model at the heart of these systems becomes exponentially more complex (Vestby and Vestby 2019). Indeed, even the most highly trained professionals would struggle to understand why these systems generate the predictions and assessments they do. In other words, as these systems improve, their opacity likewise increases.

Second, most data-driven systems are shrouded in secrecy. Some of this is due to the lack of transparency about the extent of their implementation; whether and to what extent a human is 'in the loop'; and whether, if ever, these systems are audited. A deeper concern, however, is that these systems are generally created and maintained by private companies, who retain extensive intellectual property rights over their proprietary source code. As a result, many of those subject to these systems are incapable of scrutinizing them. In *State v. Loomis*, the defendant, Eric Loomis, argued that his due process rights were violated, since he and his legal team were not granted access to the COMPAS algorithm that deemed him high risk, since this information was proprietary (*State v. Loomis* 2016). While COMPAS is a tool used in courthouses, the same issues arise with the predictive technologies at use by police departments nationwide.

Opacity in the use of data-driven policing technologies is problematic on several fronts. First, government institutions ought to be capable of publicly justifying their treatment of citizens *to* the citizens affected. But extensive technical opacity hinders their ability to do so.

Moreover, both sorts of opacity have the potential to sow distrust and skepticism toward police, which may worsen many of the problems these technologies are intended to resolve. For example, many police departments have adopted data-driven systems in response to concerns about racial bias in policing that occurs when police departments rely entirely on officers' individual judgment and discretion. In response to a critical exposé of their Intelligence Led Policing Program practices in the Tampa Bay Times, the Sheriff's Department in Pasco County, Florida wrote that their system:

> removes any opportunity for bias by removing descriptive features and focuses strictly on the criminal history of the individual, regardless of their race, gender, creed or any other identifying factors. We are surprised to see a system that is blind to anything but the criminal history of an individual is under attack, instead of being celebrated as an important step forward in our country. (Tampa Bay Times 2020)

Removing bias is of course an admirable goal. However, when the public is unable to hold these technologies accountable, public perceptions of bias and injustice are further entrenched or confirmed. This, in turn, can lead to even greater distrust of police, decreased perceptions of legitimacy, and more injustice.

7 Challenge #5 Accountability and community oversight

The democratic legitimacy of law enforcement institutions requires that there are appropriate mechanisms for holding them accountable. Accountability mechanisms foster public trust, enable the achievement of valuable institutional goals, and identify and address any problematic elements within the institution itself (Vestby and Vestby 2019). But the value of accountability is also intrinsic: part of what constitutes a legitimate institution—particularly those, like law enforcement, that wield significant authority—is that they are democratically accountable to those they are intended to serve.

The rapid introduction of, and increasing reliance upon, data-driven methods throughout law enforcement introduces a number of challenges for citizens' ability to hold accountable the institution of law enforcement and those who play a role in it. One such challenge concerns an issue mentioned in the previous section—namely, opacity. While law enforcement requires some measure of opacity or secrecy for their operations to be at all effective, secrecy threatens citizens' ability to hold law enforcement accountable.

Most of the data-driven technologies employed by police departments across the country suffer from various sorts of opacity. For one thing, many of the most high-profile technologies have been developed and maintained by private companies, which are legally afforded (and typically retain) the right not to disclose certain proprietary elements of these systems to the departments that use them, or to the general public. Police departments are often given full authority to procure these systems from vendors directly and to implement them without oversight from the public or from city commissions (Crump 2016; Joh 2017). Privatization thus poses a significant obstacle to the efforts of concerned citizens, advocacy groups, and those directly impacted by these systems to hold law enforcement accountable. Without full access to these systems, any attempts at holding law enforcement accountable will be incomplete.

One step toward addressing this issue is to require technology companies to release proprietary information to the relevant parties in certain cases. While companies have thus far largely resisted this idea, and police departments have shown little interest in insisting on it as a condition of purchase, the courts may offer some assistance. In a recent case, *New Jersey v. Pickett,* the court ruled that defendants have a right to inspect and understand the software that is used to provide evidence against them, after a defendant had previously been denied access to the proprietary DNA software that served as the basis for his arrest (Fasciale 2021). Perhaps other courts will follow suit for algorithmic technologies used in other facets of the criminal justice system, including policing.

While this case is a step in the right direction, it does not resolve the issue completely. For one thing, the right to inspection was extended only to those accused of crimes but not other interested parties, such as advocacy groups and other oversight organizations, who might play a role in helping to prevent abuses of these systems before they begin. Furthermore, courts are in general a weak mechanism for oversight of law enforcement, since the courts are constitutionally limited in the ways they can provide a check on law enforcement practices. As Erik Bakke argues, "The Fourth Amendment's protections apply only after there is an invasion of a reasonable expectation of privacy. Because predictive policing is used for surveillance, which does not require reasonable

suspicion, police can observe citizens without any court interference" (Bakke 2018, p. 143). Thus, full accountability is unlikely to be achieved through the courts alone.

It is widely thought that accountability can be improved by subjecting data-driven systems to 'algorithmic audits', in which an independent organization evaluates the system for issues like bias, unfair treatment, improper data, corrupt source codes, and other related problems. There is even legislative momentum to require that private firms regularly evaluate and remedy any flaws or biases in their algorithmic systems. In April 2019, U.S. Senators Cory Booker and Ron Wyden, along with Rep. Yvette D. Clarke introduced The Algorithmic Accountability Act, which would authorize the Federal Trade Commission to require companies to conduct regular impact assessments of certain "automated decision systems" (Booker 2019). Such a requirement should extend to those private algorithms being leased by U.S. law enforcement agencies and courts.

Joy Buolamwini's work auditing commercial facial analysis models used by IBM, Microsoft, and Megvii (Face++) provides an encouraging example of how audits can address biases embedded in AI systems. Within seven months of an initial audit demonstrating disparities in accuracy of facial analysis between men and women, and between light and dark skinned people, IBM, Microsoft, and Megvii released new versions of their software that significantly reduced those disparities (Raji and Buolamwini 2019).

Another encouraging example of algorithmic auditing in practice comes from a collaboration between ShotSpotter Technologies and NYU's Policing Project. ShotSpotter Inc. is a private technology firm whose product uses computer sensors to identify and analyze gunshots and then notify law enforcement. The Policing Project, run out of NYU's School of Law, "partners with communities and police to promote public safety through transparency, equity, and democratic engagement" ("Policing Project" 2020). ShotSpotter voluntarily approached the Policing Project to conduct an audit of potential privacy threats to bystanders whose voices might be picked up by the ShotSpotter sensors. While the audit revealed little threat of unintended voice surveillance, the Policing Project's final report on the audit happily describes that ShotSpotter "adopted nearly all of our [privacy protection] recommendations verbatim" ("Policing Project" 2020).

These stories are promising, and algorithm audits are, to be sure, an essential step in ensuring law enforcement accountability; however, several key issues remain. First, the auditing process itself requires significant transparency—that is, measures must be put in place to 'audit the auditors', or to otherwise ensure the independence of the auditing process. Without this, there will remain incentives for departments to employ only those auditors that are most favorable to them, and for auditors to provide only the most favorable results.

Moreover, even if such audits are fully independent, measures must be put in place to ensure the full unredacted reports are released without spin. Recently, a company called HireVue, which uses AI during job interviews to evaluate job candidates, hired an outside consulting firm, headed by AI expert Cathy O'Neil, to perform an audit of their software (Engler 2021). As requested, O'Neil's team focused on a specific piece of HireVue's technology; but the company used the largely positive audit to suggest that their entire system had been vindicated—a conclusion that extended far beyond the evidence from the audit. While HireVue is a public company, similar concerns about transparency in the audit process itself affect public institutions like law enforcement. To be sure,

algorithm audits are an important step, but they can also be used as a form of "accountability theater" that obscures more than it illuminates.

A further way to promote accountability is through legislative action. City commissions that oversee police departments can create citizen-led councils to act as a conduit through which the community can offer recommendations and advice to police about the impact of police work on their communities. These councils should include as part of their charge the evaluation of new technological policing systems, particularly their impact on minorities. These bodies, sometimes called "citizen advisory councils," are useful tools of communication and trust building between police and communities, but they provide limited accountability insofar as the council serves a merely advisory role; city officials or police departments can opt to ignore the recommendations. A more demanding approach would require *prior* legislative authorization by city officials of any new data-driven technologies by police ("ACLU" 2018; Ferguson 2020).

8 Conclusion

The list of challenges presented here is far from exhaustive. There are many other urgent technical, legal, and ethical challenges facing data-driven policing.

For certain of these challenges, there are clear pathways forward. For instance, as we have noted, several of the problems, such as those concerning opacity and oversight, can be partially addressed via legislation or departmental policies that require that any companies contracted to provide algorithmic technologies to police departments waive the relevant intellectual property rights in certain contexts and to submit to regular external independent audits. None of these steps is a panacea, and they each raise further issues that must be addressed as well.

Other challenges, such as those relating to bias, data selection, standards of success, and deeper questions about the social value of police require more collective deliberation. This deliberation must be informed by academics, technologists, law enforcement, and, most importantly, the lived experience of the public. Some scholars have recently considered whether addressing algorithmic bias directly in the design and programming might mitigate or eliminate these negative effects, and the initial results appear promising (Skeem and Lowenkamp 2020). There are, however, concerns with this approach; and some will find the possibility of any amount of encoded bias unacceptable.

While police departments nationwide vie for better data-driven technologies to support their mission to protect and serve their communities, advocacy groups and the broader public are calling for the discontinuation of data-driven policing programs. As we have attempted to show here, what ought to be done about data-driven policing is a complex and multi-faceted issue. It will be virtually impossible to satisfy all sides of the debate. Those who wish to continue on the path to better data-driven policing must work to address the challenges we have raised above, and many other issues that are raised by this technological shift. And those who oppose data-driven policing practices ought to be aware of and sensitive to the full range of challenges, their complexity, and the various possible pathways forward. Socially responsible data-driven policing requires collaboration between academics, technology developers, police departments, and policy-makers to confront and address

these challenges. Most importantly, it requires engagement with and oversight by affected communities, especially those who have become deeply skeptical about the enterprise of policing.

Disclosure statement: Work on this project was supported by funding from the Consortium on Trust in Media and Technology at the University of Florida and National Science Foundation Award #1917712: Artificial Intelligence and Predictive Policing: An Ethical Analysis.

Acknowledgements: Thank you to research assistants Kiana Alikhademi, Ami Asar, and Jonathan Suarez for valuable conversations and written feedback on drafts of this paper.

References

ACLU. (2018). *Community Control Over Police Surveillance #TakeCTRL*. https://communityresourcehub.org/resources/community-control-over-police-surveillance-takectrl/. Accessed 26 February 2021

Adams, M., & Rameau, M. (2016). Black Community Control Over Police. *Wisconsin Law Review*, *515*.

Bakke, E. (2018). Predictive Policing: The Argument for Public Transparency. *Annual Survey of American Law*, *74*(131), 131–172.

Barocas, S., & Selbst, A. D. (2016). Big Data's Disparate Impact. *California Law Review*, *104*(3), 671–732.

Beck, A., & Blumstein, A. (2018). Racial Disproportionality in U.S. State Prisons: Accounting for the Effects of Racial and Ethnic Differences in Criminal Involvement, Arrests, Sentencing, and Time Served. *Journal of Quantitative Criminology*, *34*, 853–883.

Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity.

Blumstein, A. (2011). Bringing Down the U.S. Prison Population. *The Prison Journal*, *91*(3).

Boba Santos, R. (2020). Predictive policing: Where's the evidence? In D. Weisburd & A. A. Braga (Eds.), *Predictive policing: Where's the evidence?* (2nd Edition.). Cambridge: Cambridge University Press.

Bobo, L. D., & Thompson, V. (2006). Unfair By Design: The War on Drugs, Race, and the Legitimacy of the Criminal Justice System. *Social Research*, *73*(2), 445–472.

Booker, C. (2019). *Booker, Wyden, Clarke Introduce Bill Requiring Companies To Target Bias In Corporate Algorithms*. https://www.booker.senate.gov/news/press/booker-wyden-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms. Accessed 15 February 2021

Brantingham, J., & Valasik, M. (2018). Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial. *Statistics and Public Policy*, *5*(1), 1–6.

Brayne, S. (2014). Surveillance and System Avoidance: Criminal Justice Contact and Institutional Attachment. *American Sociological Review*, *79*(3), 367–391.

Brayne, S. (2020). *Predict and Surveil: Data, Discretion, and the Future of Policing*. New York: Oxford University Press.

Brooks, R. R. W. (2000). Fear and Fairness in the City: Criminal Enforcement and Perceptions of Fairness in Minority Communities. *California Law Review*, *73*(1219), 1228–1229.

Burbank, K. (2021, January 14). City First In Nation To Ban Predictive Policing, Biometric Surveillance Tech. *Sfgate.com*. https://www.sfgate.com/news/bayarea/article/City-First-In-Nation-To-Ban-Predictive-Policing-15872642.php

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. *Big Data and Society*, *3*(1).

Capers, I. B. (2017). Race, Policing, and Technology. *North Carolina Law Review*, *95*(1241).

Caplan, J., Kennedy, L., & Miller, J. (2011). Risk Terrain Modeling: Brokering Criminological Theory and GIS Methods for Crime Forecasting. *Justice Quarterly*, *28*(2), 360–381.

Cohen, L. E., Kluegel, J., & Land, K. (1981). Social Inequality and Predatory Criminal Victimization: An Exposition and Test of Formal Theory. *American Socilogical Review*, *46*, 505–524.

Crump, C. (2016). Surveillance Policymaking by Procurement. *Washington Law Review*, *91*(1595).

Dumke, M., & Main, F. (2017, May 18). A look inside the watch list Chicago police fought to keep secret. *Chicago Sun-Times*. https://chicago.suntimes.com/2017/5/18/18386116/a-look-inside-the-watch-list-chicago-police-fought-to-keep-secret. Accessed 25 February 2021

Ekins, E. E. (2016, December 7). *Policing in America: Understanding Public Attitudes Toward the Police. Results from a National Survey*. Cato Institute. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2919449. Accessed 15 February 2021

Engler, A. C. (2021, January 26). Independent auditors are struggling to hold AI companies accountable. *Fast Company*. https://www.fastcompany.com/90597594/ai-algorithm-auditing-hirevue. Accessed 15 February 2021

Fasciale, P. J. A. D. State of New Jersey v. Corey Pickett. , No. A-4207-19T4 (Superior Court of New Jersey 3 February 2021). https://www.eff.org/files/2021/02/03/decision_appdiv_232021.pdf

Ferguson, A. G. (2015). Big Data and Predictive Reasonable Suspicion. *University of Pennsylvania Law Review*, *163*(2), 327–410.

Ferguson, A. G. (2017). *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: NYU Press.

Ferguson, A. G. (2020, June 26). It's time for D.C. to regulate police surveillance technology. *Washington Post*. https://www.washingtonpost.com/opinions/local-opinions/its-time-for-dc-to-regulate-police-surveillance-technology/2020/06/25/9e94feb6-b57a-11ea-aca5-ebb63d27e1ff_story.html. Accessed 26 February 2021

Hagan, J., & Albonetti, C. (1982). Race, Class, and the Perception of Criminal Injustice in America. *American Journal of Sociology*, *88*(2), 329–355.

Harcourt, B. E. (2006). *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press.

Haskins, C. (2019, February 14). Academics Confirm Major Predictive Policing Algorithm is Fundamentally Flawed. *Vice.com*. https://www.vice.com/en/article/xwbag4/academics-confirm-major-predictive-policing-algorithm-is-fundamentally-flawed

Heaven, W. D. (2020, July 17). Predictive policing algorithms are racist. They need to be dismantled. *MIT Technology Review*. https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/. Accessed 15 February 2021

Joh, E. (2017). The Undue Influence of Surveillance Technology Companies on Policing. *N.Y.U. Law Review*, *92*(102).

Kleinberg, J., Ludwig, J., Mullainathan, S., & Sunstein, C. R. (2018). Discrimination in the Age of Algorithms. *Journal of Legal Analysis*, *10*, 113–174.

Lazer, D., & Radford, J. (2017). Data Ex Machina: Introduction to Big Data. *Annual Review of Sociology*, *43*(1), 19–39.

Los Angeles Police Commission. (2019). *Review of Selected Los Angeles Police Department Data-Driven Policing Strategies*. Office of the Inspector General.

Los Angeles Police Department. (2019). *The Los Angeles Police Department Strategic Plan 2019-2021*.

McGrory, K., & Bedi, N. (2020, September 3). Targeted. *Tampa Bay Times*.

https://projects.tampabay.com/projects/2020/investigations/police-pasco-sheriff-targeted/intelligence-led-policing/. Accessed 15 February 2021

Miller, L. (2020, April 21). LAPD will end controversial program that aimed to predict where crimes would occur. *Los Angeles Times*. https://www.latimes.com/california/story/2020-04-21/lapd-ends-predictive-policing-program. Accessed 15 February 2021

Moravec, E. R. (2019, September 5). Do Algorithms Have a Place in Policing? *The Atlantic*. https://www.theatlantic.com/politics/archive/2019/09/do-algorithms-have-place-policing/596851/. Accessed 15 February 2021

Muhammad, K. G. (2019). *The Condemnation of Blackness: Race, crime, and the Making of Modern Urban America*. Harvard University Press.

Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.

Piquero, A., & Brame, B. (2008). Assessing the Race-Crime and Ethnicity-Crime Relationship in a Sample of Serious Adolescent Delinquents. *Crime & Delinquency*, *54*(3), 390–422.

Policing Project. (2020). www.policingproject.org. Accessed 26 February 2021

Raji, I. D., & Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. *AIES '19: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–435.

Ryan-Mosley, T., & Strong, J. (2020, June 5). The activist dismantling racist police algorithms. *MIT Technology Review*. https://www.technologyreview.com/2020/06/05/1002709/the-activist-dismantling-racist-police-algorithms/. Accessed 25 February 2021

Saad, L. (2020, August 5). Black Americans Want Police to Retain Local Presence. *Gallup*. https://news.gallup.com/poll/316571/black-americans-police-retain-local-presence.aspx. Accessed 15 February 2021

Selbst, A. (2017). Disparate Impact in Big Data Policing. *Georgia Law Review*, *52*(109), 109–195.

Short, M. B., D'Orsogna, M. R., Brantingham, P. J., & Tita, G. E. (2009). Measuring and Modeling Repeat and Near-Repeat Burglary Effects. *Journal of Quantitative Criminology*, *25*, 325–339.

Skeem, J., & Lowenkamp, C. (2020). Using algorithms to address trade-offs inherent in predicting recidivism. *Behavioral Sciences and the Law*, *38*(3), 259–278.

State v. Loomis. , No. 881 N.W.2d 749 (2016).

Sturgill, K. (2020, June 26). Santa Cruz becomes the first U.S. city to ban predictive policing. *Los Angeles Times*. https://www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing. Accessed 15 February 2021

Tampa Bay Times. (2020, September 3). Read the Pasco Sheriff's Office response to our investigation. *Tampa Bay Times*. https://projects.tampabay.com/projects/2020/investigations/police-pasco-sheriff-targeted/sheriffs-response/. Accessed 15 February 2021

Uchida, C. D., Swatt, M., Gamero, D., Lopez, J., Salazar, E., King, E., et al. (2012). *The Los Angeles Smart Policing Initiative: Reducing Gun-Related Violence through Operation LASER*. U.S. Department of Justice, Bureau of Justice Assistance. http://newweb.jssinc.org/wp-content/uploads/2014/11/Spotlight-on-Operation-LASER.pdf

Vestby, A., & Vestby, J. (2019). Machine Learning and the Police: Asking the Right Questions. *Policing: A Journal of Policy and Practice*.

Vitale, A. S. (2017). *The End of Policing*. Verso Books.