DISPARATE IMPACT IN BIG DATA POLICING

Andrew D. Selbst[*]

Data-driven decision systems are taking over. No institution in society seems immune from the enthusiasm that automated decision-making generates, including—and perhaps especially—the police. Police departments are increasingly deploying data mining techniques to predict, prevent, and investigate crime. But all data mining systems have the potential for adverse impacts on vulnerable communities, and predictive policing is no different. Determining individuals' threat levels by reference to commercial and social data can improperly link dark skin to higher threat levels or to greater suspicion of having committed a particular crime. Crime mapping based on historical data can lead to more arrests for nuisance crimes in neighborhoods primarily populated by people of color. These effects are an artifact of the technology itself, and will likely occur even assuming good faith on the part of the police departments using it. Meanwhile, predictive policing is sold in part as a "neutral" method to counteract unconscious biases when it is not simply

---

[*] Postdoctoral Scholar, Data & Society Research Institute; Visiting Fellow, Yale Information Society Project. J.D. 2011, University of Michigan. For helpful comments and insights on earlier drafts, I would like to thank Micah Altman, Jack Balkin, Jane Bambauer, Solon Barocas, Rabia Belt, Kiel Brennan-Marquez, William Buzbee, Danielle Citron, Julie Cohen, Jessica Eaglin, Gautam Hans, Don Herzog, Mark MacCarthy, Deirdre Mulligan, Leonard Niehoff, Paul Ohm, Christopher Slobogin, Katherine Strandburg, David Super, Michael Veale, David Vladeck, Robin West, and the participants at the Georgetown Fellows Colloquium, Georgetown Summer Faculty Workshop, Michigan Future Law Professors Workshop, and 2017 Privacy Law Scholars' Conference.

sold to cash-strapped departments as a more cost-efficient way to do policing.

The degree to which predictive policing systems have these discriminatory results is unclear to the public and to the police themselves, largely because there is no incentive in place for a department focused solely on "crime control" to spend resources asking the question. This is a problem for which existing law does not provide a solution. Finding that neither the typical constitutional modes of police regulation nor a hypothetical anti-discrimination law would provide a solution, this Article turns toward a new regulatory proposal centered on "algorithmic impact statements."

Modeled on the environmental impact statements of the National Environmental Policy Act, algorithmic impact statements would require police departments to evaluate the efficacy and potential discriminatory effects of all available choices for predictive policing technologies. The regulation would also allow the public to weigh in through a notice-and-comment process. Such a regulation would fill the knowledge gap that makes future policy discussions about the costs and benefits of predictive policing all but impossible. Being primarily procedural, it would not necessarily curtail a department determined to discriminate, but by forcing departments to consider the question and allowing society to understand the scope of the problem, it is a first step towards solving the problem and determining whether further intervention is required.

TABLE OF CONTENTS

I. INTRODUCTION

Machine-driven decision-making is everywhere. Decisions based on machine learning algorithms are supplementing or replacing human decision-making in vastly different aspects of society, including consumer finance,[1] employment,[2] housing,[3] healthcare,[4] and sentencing,[5] among others. One particularly important area of rapid adoption is predictive policing, a popular and growing method for police departments to prevent or solve crimes.[6]

Though predictive methods such as crime mapping and offender profiling are not new, predictive policing is something different, a creature of the world of Big Data. A police department engaged in predictive policing uses data mining methods to find correlations between criminal outcomes and various input data they have collected—crime locations,[7] social networks,[8] or commercial data.[9]

---

[1] *Machine-Learning Promises to Shake Up Large Swathes of Finance*, ECONOMIST (May 25, 2017), https://www.economist.com/news/finance-and-economics/21722685-fields-trading-credit-assessment-fraud-prevention-machine-learning.

[2] Tammy Wang, *How Machine Learning Will Shape the Future of Hiring*, LINKEDIN: PULSE (Mar. 8, 2017), https://www.linkedin.com/pulse/how-machine-learning-shape-future-hiring-tammy-wang.

[3] John Biggs, *Naborly Lets Landlords Screen Tenants Automagically*, TECHCRUNCH (Aug. 15, 2016), https://techcrunch.com/2016/08/15/naborly-lets-landlords-screen-tenants-automagically.

[4] Bill Siwicki, *Machine Learning 101: The Healthcare Opportunities Are Endless*, HEALTHCARE IT NEWS (Mar. 30, 2017), http://www.healthcareitnews.com/news/machine-learning-101-healthcare-opportunities-are-endless.

[5] Jason Tashea, *Courts Are Using AI to Sentence Criminals. That Must Stop Now*, WIRED (Apr. 17, 2017), https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now.

[6] *See* Ellen Huet, *Server and Protect*, FORBES, Mar. 2, 2015, at 46 ("In a 2012 survey of almost 200 police agencies 70% said they planned to implement or increase use of predictive policing technology in the next two to five years.").

[7] Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 277 (2012).

[8] *See* Jennifer Bachner, *Predictive Policing: Preventing Crime with Data and Analytics*, IBM CTR. FOR THE BUS. OF GOV'T 22 (2013), http://www.businessofgovernment.org/sites/default/files/Predictive%20Policing.pdf ("Through [social network analysis], police can identify individuals that are central to criminal organizations such as gangs and drug networks, and develop effective interdiction strategies."); Jennifer A. Johnson et al., *Social Network Analysis: A Systematic Approach for Investigating*, FBI L. ENFORCEMENT BULL. (Mar. 5, 2013), https://leb.fbi.gov/2013/march/social-network-analysis-a-systematic-approach-for-investigating (describing how law enforcement employs "social network analysis" to "discover, analyze, and visualize the social networks of criminal suspects").

Predictive policing is the melding of "information technology . . ., criminology theory, [and] predictive algorithms."[10] Simply put, it is "the use of data and analytics to predict crime."[11]

Despite its growing popularity, predictive policing is in its relative infancy and is still mostly hype.[12] Current prediction is akin to early weather forecasting,[13] and, like Big Data approaches in other sectors,[14] mixed evidence exists about its effectiveness. Cities such as Los Angeles, Atlanta, Santa Cruz, and Seattle have enlisted the predictive policing software company PredPol to predict where property crimes will occur.[15] Santa Cruz reportedly "saw burglaries drop by 11% and robberies by 27% in the first year of using [PredPol's] software."[16] Similarly, Chicago's Strategic Subject List—or "heat list"—of people most likely to be involved in a shooting had, as of mid-2016, predicted more than 70% of the

---

[9] WALTER L. PERRY ET AL., PREDICTIVE POLICING: THE ROLE OF CRIME FORECASTING IN LAW ENFORCEMENT OPERATIONS 13 (2013), https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf ("Free and commercial data sets are available for use with crime data; examples of useful analytic additions include data on businesses, infrastructure, and demographics.").

[10] Ferguson, *supra* note 7, at 265.

[11] Bachner, *supra* note 8, at 6; *see also* CRAIG D. UCHIDA, NAT'L INST. OF JUSTICE, NO. NCJ 230404, A NATIONAL DISCUSSION ON PREDICTIVE POLICING: DEFINING OUR TERMS AND MAPPING SUCCESSFUL IMPLEMENTATION STRATEGIES 1 (2009), https://www.ncjrs.gov/pdffiles1/nij/grants/230404.pdf ("Predictive policing refers to any policing strategy or tactic that develops and uses information and advanced analysis to inform forward-thinking crime prevention." (emphasis omitted)).

[12] Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1125, 1143 (2017); *see also* PERRY ET AL., *supra* note 9, at 115 ("[P]redictive policing has been so hyped that the reality cannot live up to the hyperbole."); Darwin Bond-Graham & Ali Winston, *All Tomorrow's Crime: The Future of Policing Looks a Lot Like Good Branding*, S.F. WEEKLY (Oct. 30, 2013), http://www.sfweekly.com/news/all-tomorrows-crime s-the-future-of-policing-looks-a-lot-like-good-branding/ (noting the media hype of PredPol's predictive policing software).

[13] *See* Ferguson, *supra* note 12, at 1143–44 (explaining that data, like early weather forecasting, can provide localized results "with a significant degree of variability and fallibility" but that "the move to objective, data-driven computer models signals an improvement from subjective instincts or traditional guesses about the weather"); Lawrence W. Sherman, *The Rise of Evidence-Based Policing: Targeting, Testing, and Tracking*, 42 CRIME & JUST. 377, 427 (2013) ("Like weather forecasting, individual forecasting or crime risks uses the recent growth in supercomputers to find highly specific combinations of predictors that raise the odds of fairly rare events occurring.").

[14] *See* Simon Moss, *Big Data: New Oil or Snake Oil?*, WIRED, https://www.wired.com/insights/2014/10/big-data-new-oil-or-snake-oil/ (last visited Sept. 25, 2017).

[15] *See, e.g.*, Bond-Graham & Winston, *supra* note 12.

[16] Huet, *supra* note 6, at 46.

people shot in the city, according to the police.[17] But two rigorous academic evaluations of predictive policing experiments, one in Chicago and another in Shreveport, have shown no benefit over traditional policing.[18] A great deal more study is required to measure both predictive policing's benefits and its downsides.

One potential downside is clear. As Solon Barocas and I observed in an earlier work, if its users are not careful, "data mining can reproduce existing patterns of discrimination, inherit the prejudice of prior decision makers, or simply reflect the widespread biases that persist in society."[19] In August 2016, seventeen civil rights organizations released a joint statement on the civil rights concerns of predictive policing, emphasizing the possibility of racist outcomes, as well as the lack of transparency, public debate, and attention to community needs.[20] The way police are adopting and using these technologies means more people of color are arrested, jailed, or physically harmed by police, while the needs of communities being policed are ignored.

Like other sectors' use of data mining,[21] predictive policing is sold in part as a way to counteract the conscious or unconscious

---

[17] Monica Davey, *Chicago Police Try to Predict Who May Shoot or Be Shot*, N.Y. TIMES (May 23, 2016), https://www.nytimes.com/2016/05/24/us/armed-with-data-chicago-police-try-to-predict-who-may-shoot-or-be-shot.html.

[18] PRISCILLIA HUNT ET AL., EVALUATION OF THE SHREVEPORT PREDICTIVE POLICING EXPERIMENT 33 (2014), https://www.rand.org/content/dam/rand/pubs/research_reports/RR500/RR531/RAND_RR531.pdf (finding no statistically significant decrease in property crime as a result of a predictive policing effort in Shreveport, Louisiana); Jessica Saunders et al., *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. EXPERIMENTAL CRIMINOLOGY 347, 366 (2016); *see also* DAVID ROBINSON & LOGAN KOEPKE, STUCK IN A PATTERN: EARLY EVIDENCE ON "PREDICTIVE POLICING" AND CIVIL RIGHTS 7–8 (2016), http://centerformediajustice.org/wp-content/uploads/2016/08/Upturn_-_Stuck_In_a_Pattern_v.1.01.pdf (noting that the foregoing are the only two scholarly studies produced by authors without a financial or reputational interest in the outcome).

[19] *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 674 (2016). Throughout this Article I will focus on racial discrimination, primarily because that is the focus of the broader discrimination discussion as it pertains to policing. The arguments about data and discrimination apply equally well to other classes of vulnerable populations based on gender, gender identity, sexual orientation, as well as non-legally protected classes such as social class. *See generally id.*

[20] *See* THE LEADERSHIP CONFERENCE ON CIVIL & HUMAN RIGHTS ET AL., PREDICTIVE POLICING TODAY: A SHARED STATEMENT OF CIVIL RIGHTS CONCERNS (Aug. 31, 2016) [hereinafter STATEMENT OF CIVIL RIGHTS GROUPS], http://civilrightsdocs.info/pdf/FINAL_JoIntStatementPredictivePolicing.pdf.

[21] *See, e.g.*, Daniel Castro, *Data Detractors Are Wrong: The Rise of Algorithms Is a Cause for Hope and Optimism*, CTR. FOR DATA INNOVATION (Oct. 25, 2016), https://www.datainnov

prejudices of human decision-makers—in this case the police.[22] And it has the potential to do so. But while most predictive policing systems will not consider race expressly, express consideration of race is not necessary for data mining to have a disproportionate racial impact.[23] A data mining system incorporates a series of man-made decisions that can create or exacerbate discriminatory outcomes, independent of any intent to do so.[24]

While policing is just one of many aspects of society being upended by machine learning, and potentially exacerbating disparate impact in a hidden way as a result, it is a particularly useful case study because of how little our legal system is set up to regulate it. Traditionally, the Fourth Amendment has been seen as the primary means by which Americans can regulate police.[25] To those familiar with the relationship between the Fourth Amendment and race, this does not inspire hope. Case law has largely removed claims of racial discrimination from the purview of the Fourth Amendment,[26] and it will not provide a solution here.

---

ation.org/2016/10/data-detractors-are-wrong-the-rise-of-algorithms-is-a-cause-for-hope-and-optimism/ (arguing that many companies use big data to counteract bias in hiring and other corporate activities).

[22] *See* PERRY ET AL., *supra* note 9, at 1 ("Forecasting is considered objective, scientific, reproducible, and free from individual bias and error."); Joshua Brustein, *This Guy Trains Computers to Find Future Criminals*, BLOOMBERG (July 18, 2016), https://www.bloomberg.com/features/2016-richard-berk-future-crime/ ("Richard Berk says his algorithms take the bias out of criminal justice."). *But see* Seeta Peña Gangadharan, *Predictive Algorithms Are Not Inherently Unbiased*, N.Y. TIMES: ROOM FOR DEBATE (Nov. 19, 2015), https://www.nytimes.com/roomfordebate/2015/11/18/can-predictive-policing-be-ethical-and-effective/predictive-algorithms-are-not-inherently-unbiased ("A misguided belief in the objectivity and neutrality of predictive technologies permeates every step of the process—from the software developers . . . to frontline officers responding to computer recommendations of crime 'hot spots,' to the local administrators monitoring police department performance.").

[23] *See* Moritz Hardt, *Approaching Fairness in Machine Learning*, MOODY RD (Sept. 6, 2016), http://blog.mrtz.org/2016/09/06/approaching-fairness.html ("Historically, the naive approach to fairness has been to assert that the algorithm simply doesn't look at *protected attributes* such as race, color, religion, gender, disability, or family status. So, how could it discriminate? This idea of *fairness through blindness*, however, fails due to the existence of *redundant encodings*. There are almost always ways of predicting unknown protected attributes from other seemingly innocuous features.").

[24] *See* Barocas & Selbst, *supra* note 19, at 673–74.

[25] *See, e.g.*, Christopher Slobogin, *Policing as Administration*, 165 U. PA. L. REV. 91, 96 (2016) (referring to "the understandable belief that the Fourth Amendment, as a practical matter, has preempted the field of police regulation").

[26] *See, e.g.*, Whren v. United States, 517 U.S. 806, 813 (1996).

As Andrew Ferguson has observed, the Fourth Amendment's reasonable suspicion requirement is inherently a "small data doctrine," rendering it impotent in even its primary uses when it comes to data mining.[27]

If the existing legal constraints cannot address the issues, new legal strategies are needed. Accordingly, this Article joins the growing call for ex ante regulation of police.[28] Generally, there are good arguments for administrative or legislative regulation of police rather than constitutional. As Barry Friedman and Maria Ponomarenko argue, the need for democratic accountability is a strong normative argument in favor of ex ante, transparent regulations.[29] Christopher Slobogin argues that notice-and-comment procedures are needed for police because the Fourth Amendment is not cutting it.[30] He argues that current case law forks between a toothless "special needs" doctrine that permits police to do essentially anything if they say it is being done for non-criminal reasons (even though administrative searches often seek criminal activity in practice),[31] and an individualized search doctrine that judges are unwilling or unable to use to examine large-scale programmatic practices.[32] Daphna Renan similarly

---

[27] Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 338 (2015).

[28] Andrew Manuel Crespo, *Systemic Facts: Toward Institutional Awareness in Criminal Courts*, 129 HARV. L. REV. 2049, 2051 (2016); Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1833 (2015); Slobogin, *supra* note 25, at 95. This push actually originated decades ago, even as the Warren Court was enshrining the Fourth Amendment's place as the primary legal constraint on the police, but the original movement never took hold. *See* David A Slansky, *Quasi-Affirmative Rights in Constitutional Criminal Procedure*, 88 VA. L. REV. 1229, 1272–73 (2002) (discussing early scholarship of the mid-1970s that argued for greater police rulemaking and judicial oversight).

[29] Friedman & Ponomarenko, *supra* note 28, at 1889. Friedman and Ponomarenko note that the use of "technologies that could not have been envisioned" a long time ago, when police received general legislative grants to investigate crime, is a circumstance "in which public rulemaking seems particularly essential" from a democratic accountability standpoint. *Id.* at 1884.

[30] *See* Slobogin, *supra* note 25, at 151 (summarizing the inherent problems in Fourth Amendment jurisprudence while arguing that "[a] regulatory regime based on administrative law principles would hold law enforcement agencies more accountable").

[31] *Id.* at 109–10; Christopher Slobogin, *Government Dragnets*, 73 LAW & CONTEMP. PROBS. 107, 130 (2010) (discussing how noncriminal laws often lead to "pretextual dragnets"); *see also* Eve Brensike Primus, *Disentangling Administrative Searches*, 111 COLUM. L. REV. 254, 259 (2011).

[32] Tracey L. Meares, *Programming Errors: Understanding the Constitutionality of Stop-and-Frisk as a Program, Not an Incident*, 82 U. CHI. L. REV. 159, 164 (2015); Daphna

argues that administrative law can bolster a Fourth Amendment doctrine ill-suited to the programmatic nature of modern-day policing.[33]

Regarding predictive policing specifically, society lacks basic knowledge and transparency about both the technology's efficacy and its effects on vulnerable populations.[34] Thus, this Article proposes a regulatory solution designed to fill this knowledge gap—to make the police do their homework and show it to the public before buying or building these technologies.

The Article proceeds in three Parts. Part I offers a sketch of predictive policing from a technical perspective: how it works, how it is used, and how the results will impact communities of color. Predictive policing systems rely heavily on past crime data,[35] which will inevitably reproduce past biases existing in the collection of such data. But there are several other hidden mechanisms by which data mining systems create or exacerbate disparate impact, and this Part will discuss those as well.

Part II examines the failures of various standing legal strategies. Two possibilities are considered within existing Fourth Amendment law: direct application of disparate impact doctrine, and the use of the individualization requirement as a substitute. A judicial commitment to a colorblind Fourth Amendment renders the adoption of disparate impact doctrine unlikely. And although the Fourth Amendment can address strict racial profiling through the requirement of individualized suspicion, that effect does not translate to a predictive policing system that only incidentally relies on racial proxies. Drawing on prior work in Title VII jurisprudence, this Part ends with a discussion of why, even if disparate impact doctrine were incorporated into Fourth Amendment law or adopted by statute, it would still not address the issue.

Part III introduces "algorithmic impact statements." Modeled on the environmental impact statements of the National

---

Renan, *The Fourth Amendment as Administrative Governance*, 68 STAN L. REV. 1039, 1042–43 (2016).

[33] *See generally* Renan, *supra* note 32.

[34] *See* Ferguson, *supra* note 12, at 1165–68; ROBINSON & KOEPKE, *supra* note 18, at 9–10.

[35] ROBINSON & KOEPKE, *supra* note 18, at 3.

Environmental Policy Act,[36] this proposed regulation would mandate that, before adopting the new technology, police consider and publicly detail the predicted efficacy of and disparate impact resulting from their choice of technology and all reasonable alternatives. While this proposal will not rectify all that is deficient in the failed regimes discussed in Part II, it is not intended to. Rather, impact statements are designed to force consideration of the problem at an early stage, and to document the process so that the public can learn what is at stake, perhaps as a precursor to further regulation. The primary problem is that no one, including the police using the technology, yet knows what the results of its use actually are.

Fundamentally, this is more than a policing paper. The Article is about society's aggressive and uninformed move toward reliance on machine learning technologies; policing is but one example of many. The same lack of transparency and democratic buy-in that exists in predictive policing also appears in technologies used in many other sectors. Ultimately, impact assessments can be a tool to fill in knowledge that is currently lacking—knowledge necessary to even determine how future regulation should proceed. While this Article is about the police, impact statements can be used more broadly still.

II. THE DISCRIMINATORY EFFECTS OF PREDICTIVE POLICING

Police act with incredible discretion. They choose where to focus their attention, who to arrest, and when to use force. They make many choices every day regarding who is a suspect and who appears to be a criminal.[37] Examined in the aggregate, all of those choices exhibit disproportionate impacts on poor people and people of color.[38] This is the result of bias built into policing as an institution, as well as unconscious biases of individual police

---

[36] 42 U.S.C. § 4321 (2012).

[37] PERRY ET AL., *supra* note 9, at 41–49.

[38] *See, e.g.*, Tracey Maclin, *Race and the Fourth Amendment*, 51 VAND. L. REV. 331, 344–54 (1998) (detailing data demonstrating the effect of discriminatory police practices on people of color); Kia Makarechi, *What the Data Really Says About Police and Racial Bias*, VANITY FAIR (July 14, 2016), http://www.vanityfair.com/news/2016/07/data-police-racial-bias (summarizing research finding racial bias in policing).

officers.[39] Thus, where police use predictive policing technology, the purpose is not only to detect hidden patterns, but also to inject a "neutral," data-driven tool into the process to prevent unconscious police biases from entering the equation.[40] Predictive policing promises both to provide auditable methods that will prevent invidious intentional discrimination and to mitigate the unconscious biases attending police officers' daily choices.

But at the moment, such a promise amounts to little more than a useful sales tactic. Data mining is likely to introduce new discrimination or to reproduce and exacerbate the existing discrimination in society due to various design choices that are necessary to any data mining system.[41] Over the last few years, examples of such discriminatory outcomes have appeared in the news repeatedly. Google's AdWords unintentionally linked "black-sounding" names to criminal records.[42] Amazon unintentionally excluded minority neighborhoods from same-day delivery services.[43] Risk assessment scores used in criminal sentencing overestimate black recidivism and underestimate white recidivism.[44]

---

[39] *See* Adam Benforado, *Frames of Injustice: The Bias We Overlook*, 85 IND. L.J. 1333, 1367 (2010) (discussing implicit biases in the criminal justice system); L. Song Richardson, *Arrest Efficiency and the Fourth Amendment*, 95 MINN. L. REV. 2035, 2039–40 (2011) (explaining how both conscious and unconscious biases influence police officers' actions); Anthony C. Thompson, *Stopping the Usual Suspects: Race and the Fourth Amendment*, 74 N.Y.U. L. REV. 956, 984–91 (1999) (explaining social science data outlining how racial schemas influence police decision-making).

[40] *See* Mara Hvistendahl, *Can 'Predictive Policing' Prevent Crime Before It Happens?*, SCIENCE (Sept. 28, 2016), http://www.sciencemag.org/news/2016/09/can-predictive-policing-prevent-crime-it-happens ("[T]he algorithms can help bring down crime rates while also reducing bias in policing, their creators say. They replace more basic trendspotting and gut feelings about where crimes will happen and who will commit them with ostensibly objective analysis.").

[41] *See* Barocas & Selbst, *supra* note 19, at 677–93.

[42] *Racism is Poisoning Online Ad Delivery, Says Harvard Professor*, MIT TECH. REV. (Feb. 4, 2013), https://www.technologyreview.com/s/510646/racism-is-poisoning-online-ad-delivery-says-harvard-professor/.

[43] David Ingold & Spencer Soper, *Amazon Doesn't Consider the Race of Its Customers. Should It?*, BLOOMBERG (Apr. 21, 2016), http://www.bloomberg.com/graphics/2016-amazon-same-day/ (noting that "[i]n six major same-day delivery cites . . . the service area excludes predominantly black ZIP codes").

[44] *See* Julia Angwin et al., *Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; *see also* Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 STAN. L. REV. 803, 821–36 (2014); Danielle Citron, *(Un)fairness of Risk*

Researchers are finding evidence of racially biased outcomes in predictive policing as well. In 2016, Kristian Lum and William Isaac of the Human Rights Data Analysis Group published a case study analyzing the effects of a predictive policing algorithm if it were applied in Oakland to police drug crime.[45] They found that "the police data appear to disproportionately represent crimes committed in areas with higher populations of non-white and low-income residents."[46] Taking a more theoretical approach, computer scientists Danielle Ensign et al. have found that, without modifications to account for biases, the most minor difference in crime rates between two jurisdictions will lead to "runaway feedback loops" that cause police to focus entirely on a single jurisdiction, out of proportion with crime rates.[47]

In addition to evidence of predictive policing bias, risk assessment scores can serve as a useful parallel because they have some overlapping goals with predictive policing.[48] Specifically, both methods seek to assess the likelihood that a particular person will commit another crime in the future. However, risk assessments have been shown to be highly problematic. ProPublica reported that one company's software—the most popular one used—"was particularly likely to falsely flag black defendants as future criminals, wrongly labeling them this way at almost twice the rate as white defendants," and also routinely "mislabeled [white defendants] as low risk more often than black defendants."[49] While the journalistic attention is more recent,

---

Scores in Criminal Sentencing, FORBES (July 13, 2016), https://www.forbes.com/sites/daniellecitron/2016/07/13/unfairness-of-risk-scores-in-criminal-sentencing/#4cae4bbd4ad2.

45 Kristian Lum & William Isaac, *To Predict and Serve*?, 13 SIGNIFICANCE, Oct. 2016, at 14 (2016), http://onlinelibrary.wiley.com/doi/10.1111/j.1740-9713.2016.00960.x/epdf.

46 *Id.* at 17.

47 Danielle Ensign et al., *Runaway Feedback Loops in Predictive Policing* (July 4, 2017) (unpublished manuscript), https://arxiv.org/pdf/1706.09847v2.pdf.

48 More specifically, with person-based predictive policing. *See* discussion *infra* Part II.B.*2*.

49 Angwin et al., *supra* note 44. Northpointe, the software developer that was the subject of the ProPublica study, has responded that this imbalance is merely a reflection of differing base rates of recidivism between races, and that we should focus on the rates of true positives rather than false positives. WILLIAM DIETRICH ET AL., COMPAS RISK SCALES: DEMONSTRATING ACCURACY EQUITY AND PREDICTIVE PARITY 1–2 (2016), http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf. There are two problems with this response. First, their claim is unverifiable because we do not have reliable crime data—it is therefore not clear that we even know what the base rates are.

several scholars have discussed the potential for discrimination in risk assessment in recent years.[50] It should, therefore, not be surprising that predictive policing can reproduce or exacerbate the discrimination present in traditional policing.[51] It is likely that the discrimination in predictive policing has not been directly observed primarily because such assessments either do not exist or are proprietary.[52]

Predictive policing's limited successes may thus be overshadowed by uncertainty about its risks. Its advocates are aware that an overreliance on technology can "distract attention from the harder and more important parts of [the] process, the

See infra text accompanying notes 291–92. Second, Northpointe incorrectly presupposes that demographic parity is the only relevant measure of fairness, when in fact various notions of fairness are in play. *See, e.g.*, Moritz Hardt et al., *Equality of Opportunity in Supervised Learning*, 29 ADVANCES IN NEURAL INFO. PROCESSING SYSTEMS 3315 (2016), https://papers.nips.cc/paper/6374-equality-of-opportunity-in-supervised-learning.pdf.

[50] *See, e.g.*, Kelly Hannah-Moffat, *The Uncertainties of Risk Assessment: Partiality, Transparency, and Just Decisions*, 27 FED. SENT'G REP. 244, 244 (2015) (arguing that evidence-based risk assessments are not more accurate predictors of recidivism and do not eliminate discrimination in sentencing); Bernard E. Harcourt, *Risk as a Proxy for Race: The Dangers of Risk Assessment*, 27 FED. SENT'G REP. 237, 240 (2015) (explaining that "relying on prediction instruments . . . will surely aggravate what is already an unacceptable racial disproportionality in our prisons"); Cecelia Klingele, *The Promises and Perils of Evidence-Based Corrections*, 91 NOTRE DAME L. REV. 537, 538–41 (2015) (examining the unintended consequences of "evidence-based practices" on incarceration); Sonja B. Starr, *The New Profiling: Why Punishing Based on Poverty and Identity Is Unconstitutional and Wrong*, 27 FED. SENT'G REP. 229, 229–30 (2015) (arguing that certain variables used in risk assessments for sentencing are unconstitutional).

[51] *See* Ferguson, *supra* note 12, at 1148 ("The assumptions behind predictive technologies are affected by unseen influences that may have unintended and discriminatory consequences."); Hannah-Moffat, *supra* note 50, at 245 ("Although risk assessment tools are characterized as objective, their scoring methods and structures actually obscure the subjective and arbitrary nature of the questions and judgments they contain.").

[52] *See infra* Part IV. Despite the lack of observed phenomena, several scholars have begun to address the relationship between predictive policing and race. *See generally* ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT (2017); BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 147–60 (2007) (criticizing actuarial methods in policing because they tend to skew the results against the higher-offending group); Shima Baradaran, *Race, Prediction, and Discretion*, 81 GEO. WASH. L. REV. 157, 176–77 (2013) (noting that "[s]ome scholars claim that prediction is partially to blame for racial bias in criminal justice"); Ric Simmons, *Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*, 2016 MICH. ST. L. REV. 947, 969–75 (detailing racial biases in predictive algorithms); Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1528–29 (arguing that "automated prediction can lead to illegal discrimination" based on personal attributes such as race).

parts that rely on imagination and judgment."[53] Though "humans remain—by far—the most important elements in the process,"[54] automation bias can set in and convince the human operators that the machine knows better than they do.[55] This is particularly disconcerting with respect to hidden, systemic biases in the data.[56]

Given the history of racially discriminatory policing, it is especially important that police understand their tools' capacity for discriminatory outcomes and vigilantly guard against them.[57] Predictive policing systems operate in different ways, depending on the type of data they collect and what they seek to achieve. This section explains how predictive policing works at a technical level, and why that will result in a disparate impact on communities of color.

A. SOME CLARIFICATION ON TERMINOLOGY

As an initial matter, it will be useful to define some terms. The words "discrimination," "fairness," and "bias" evoke a family of related concepts, and their use in this Article will benefit from disambiguation. Merriam-Webster offers three definitions of "discrimination," each differently applicable in the data mining context: (1) "the practice of unfairly treating a person or group of people differently from other people or groups of people"; (2) "the ability to recognize the difference between things that are of good quality and those that are not"; and (3) "the ability to understand that one thing is different from another thing."[58] Data mining is the process of finding patterns among different people or outcomes

---

[53] DAVID ALAN SKLANSKY, *The Persistent Pull of Police Professionalism*, NEW PERSPECTIVES IN POLICING 9–10 (Mar. 2011).

[54] PERRY ET AL., *supra* note 9, at 117.

[55] *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1271–72 (2008); *see also* Cathy O'Neil, *The Ethical Data Scientist*, SLATE (Feb. 4, 2016), http://www.slate.com/articles/technology/future_tense/2016/02/how_to_bring_better_ethics_to_data_science.html ("[P]eople have too much trust in data to be intrinsically objective . . . .").

[56] *See* PERRY ET AL., *supra* note 9, at 116, 122 (noting that "biases in the inputs will skew the predictions" and that "it is important to understand how the data are collected because they may have systematic biases").

[57] *See* HARCOURT, *supra* note 52, at 169–70 (describing the social costs that will result if predictive policing tools increase actual or perceived discriminatory outcomes).

[58] *Discrimination*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/discrimination (last visited Sept. 17, 2017).

to determine what aspects make them similar or different.[59] Thus, the ultimate goal of the data miner is to build a system that can discriminate in the third sense. And to the extent the data mining system is used in a ranking scheme, the second definition is just as clearly implicated. It is the first definition, however, that is the concern of both this Article and the broader algorithmic accountability movement.[60]

As the first definition suggests, discrimination in the legal sense is closely tied to the concept of "fairness." In fact, so as to avoid ambiguity, the computer science community that concerns itself with these same issues will often refer to "fairness" rather than "discrimination."[61] But further disambiguation is required. It is possible for either a system or person to "unfairly treat[ ] a person or group of people." If the *system* treats someone unfairly, it is not necessarily because any person intended such a result, therefore, systemic discrimination is measured by its effect. In American anti-discrimination jurisprudence, this maps onto "disparate impact" doctrine, albeit imperfectly.[62] Contrarily, if a *person* treats someone unfairly on account of membership in a protected class, such action relates to intent and maps onto disparate treatment doctrine.[63] A third conceptual category exists—that of "classificatory harms"—but this separate concept is contained doctrinally within disparate treatment, and it is not

---

[59] *See, e.g.*, Liane Colonna, *A Taxonomy and Classification of Data Mining*, 16 SMU SCI. & TECH. L. REV. 309, 310 (2013) (describing data mining generally); Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 GA. L. REV. 1, 13–14 (2005) (same).

[60] *See, e.g.*, *More Accountability for Big-Data Algorithms*, NATURE (Sept. 21, 2016), http://www.nature.com/news/more-accountability-for-big-data-algorithms-1.20653 ("Fortunately, a strong movement for greater 'algorithmic accountability' is now under way in academia and, to their credit, parts of the tech industry such as Google and Microsoft.").

[61] *See, e.g.*, Cynthia Dwork et al., *Fairness Through Awareness*, 3 PROC. INNOVATIONS THEORETICAL COMPUTER SCI. CONF. 214 (2012). As of 2015, there is an annual conference dedicated to questions of fairness in machine learning, called the Conference on Fairness, Accountability and Transparency (FAT*) (formerly "Fairness, Accountability and Transparency in Machine Learning (FAT/ML)"). *See* CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAT*), http://www.fatconference.org/ (last visited Oct. 25, 2017).

[62] *See* Peter E. Mahoney, *The End(s) of Disparate Impact: Doctrinal Reconstruction, Fair Housing and Lending Law, and the Antidiscrimination Principle*, 47 EMORY L.J. 409, 411 (1998) (explaining that the same standard is "known variously as 'the effects test,' 'discriminatory effects,' or (most commonly), 'disparate impact' ").

[63] *See* Richard Primus, *The Future of Disparate Impact*, 108 MICH. L. REV. 1341, 1351 & n.56 (2010).

universally recognized as discrimination without more.[64] For all three concepts, the law only recognizes the policy as "discrimination" if it is directed at a protected class, such as race, gender, age, disability, or in some cases, sexual orientation.[65]

This Article addresses only disparate impact. This is because disparate impact is the most likely form of discriminatory harm to result from data mining.[66] With respect to data mining, a classificatory harm means the use of protected class identifiers as inputs to predictive policing models.[67] This is not currently done—presumably for fear of running afoul of equal protection doctrine.[68] Otherwise, if police want to intentionally discriminate, there are easier ways of getting away with it than hiding it in the data. This Article also uses the phrase "disparate impact" as shorthand for the general concept of measuring discrimination by effect rather than for the doctrinal rule. This is because, in the policing context, there is no extant doctrinal hook for disparate impact,[69] so there is no doctrinal referent. Because it only addresses disparate impact, this Article will refer interchangeably to discrimination and disparate impact, such as in the phrase "discriminatory data mining."

This Article also uses the word "bias" in several instances. The word "bias" is often used to mean "prejudice" or "intentional discrimination," as when describing a person as "biased."[70] That is not the case here. Instead, the word has two meanings in this Article. When referring to a bias in a model, it is used in the statistical sense to mean a factor that changes the result in a way

---

[64] *Id.*

[65] *See, e.g.*, 42 U.S.C. § 2000e-2(a) (2012) (delineating Title VII's protected classes as "race, color, religion, sex, or national origin").

[66] Barocas & Selbst, *supra* note 19, at 693.

[67] *See id.* at 695.

[68] However, there is good evidence that use of protected class identifiers could actually help rid the machine learning models of the disparate impact described in the next section, and may, in fact, be necessary. *See, e.g.*, Indrė Žliobaitė & Bart Custers, *Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-Driven Decision Models*, 24 ARTIFICIAL INTELLIGENCE L. 183, 185 (2016).

[69] Policing is primarily regulated through the Constitution, and equal protection doctrine does not recognize disparate impact. *See* Washington v. Davis, 426 U.S. 229, 242 (1976).

[70] *Bias*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/bias (last visited Sept. 17, 2017).

that is unaccounted for.[71] This Article also refers to people's "unconscious" or "implicit" biases, which refer to the implicit associations that people make without realizing they do so—a subject of a great deal of psychological and legal research at the end of the last century.[72]

Finally, there is one part of the discrimination discussion that requires disambiguation, not as a matter of terminology, but as an actual conceptual disagreement over whether a result should be considered discrimination. This occurs when a data mining model makes decisions based on proxies for protected classes and, in the process, rediscovers inequalities already present in society.[73] The disagreement over whether this is discrimination at all is explored later,[74] and that discussion points to the important distinction between competing conceptions of "fairness" as well.[75]

B. HOW PREDICTIVE POLICING DISCRIMINATES

The use of predictions in policing is not a new construct. Crime mapping, which allows the police to allocate more resources to where crime is more likely to occur, has been around for a very long time.[76] Police used to plot crime on a map by hand to see if hot spots emerged.[77] Offender profiling, in which police examine psychological and environmental factors to predict an unknown suspect's identity or to anticipate the next crime, is another form of prediction that has been around for ages.[78] Despite being

---

[71] *See id.* ("(d)(1): deviation of the expected value of a statistical estimate from the quantity it estimates (2): systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over others.").

[72] *See infra* notes 201–03 and accompanying text.

[73] *See* Barocas & Selbst, *supra* note 19, at 691–92.

[74] *See* discussion *infra* Parts II.C, IV.B.*2*.

[75] Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, PROC. INNOVATIONS THEORETICAL COMPUTER SCI. (2017), https://arxiv.org/pdf/1609.05807v1.pdf.

[76] *See* Andrew Guthrie Ferguson, *Crime Mapping and the Fourth Amendment: Redrawing "High-Crime Areas,"* 63 HASTINGS L.J. 179, 184 (2011).

[77] Andrew Ferguson points readers to a story in which a police chief in Lincoln, Nebraska put pins in a map during the Teddy Roosevelt administration. *Id.* at 184 n.24 (citing SPENCER CHAINEY & JERRY RATCLIFFE, GIS AND CRIME MAPPING 8 (2005)).

[78] *See* PETER B. AINSWORTH, OFFENDER PROFILING AND CRIME ANALYSIS 7–15 (2001) (noting that offender profiling "generally refers to the process of using all the available information about a crime, a crime scene, and a victim, in order to compose a profile of the (as yet) unknown perpetrator"); BRENT E. TURVEY, CRIMINAL PROFILING: AN INTRODUCTION

backward-looking in time, offender profiling is similar to prediction because rather than asking whether a given suspect committed the crime, the police create a model to guess the traits of a person who would likely have committed the crime.[79]

Nonetheless, the concept of "predictive policing" is something new. The phrase typically refers to policing methods that incorporate data mining.[80] Predictive policing is the logical extension of crime mapping and offender profiling to a world with more available data and processing power.[81] The fundamental premise behind profiling is that a good portion of crime occurs in predictable patterns, and if police can root out those patterns, they can either prevent crime or catch the criminals.[82] According to advocates, predictive policing performs essentially the same operation as criminal profiling, but because it uses more data and computers to find the patterns, it is both more accurate and more reliable.[83]

Data mining is the use of machine learning techniques to find useful patterns and relationships in data.[84] It works by exposing a machine learning algorithm to examples of cases of interest with known outcomes.[85] The computer then builds a predictive model—a set of correlations that determine which related attributes can serve as useful proxies for an otherwise unobservable outcome. Once those attributes are discovered, the computer compares new subjects' traits to those observed attributes to make a prediction about the unobservable outcome.[86]

TO BEHAVIORAL EVIDENCE ANALYSIS 38 (4th ed. 2012) ("Criminal profiling has a legal history that can be traced back to the blood libeling of Jews in Rome, 38 CE.").

[79] Anticipating future acts specifically is also called "forecasting," but the use of statistical techniques is common throughout predictive policing. PERRY ET AL., *supra* note 9, at 1.

[80] *See* Bachner, *supra* note 8, at 9.

[81] *See* Ferguson*, supra* note 7, at 270–72.

[82] *See id*. at 271–284.

[83] *See* Boonsri Dickinson, *In Philadelphia, Prediction and Probability in Crime Patterns*, ZDNET (Aug. 19, 2010), http://www.zdnet.com/article/in-philadelphia-prediction-and-probability-in-crime-patterns/ (discussing the algorithms' improvements on techniques in use since the 1920s that lacked support from fundamental science).

[84] *See, e.g.*, Usama Fayyad, *The Digital Physics of Data Mining*, 44 COMM. ACM 62, 62 (Mar. 2001) (describing data mining generally).

[85] *See* Pedro Domingos, *A Few Useful Things to Know About Machine Learning*, 55 COMM. ACM 78, 78–80 (Oct. 2012).

[86] *See* Barocas & Selbst, *supra* note 19, at 678.

In predictive policing, the observed attributes come from data that the police mine from various sources. There are several different approaches to predictive policing. Some primarily use data about past criminal activity, such as crime locations and arrest records, but others incorporate many other types of data. These companies sometimes purchase tools "largely developed by and for the commercial world,"[87] as well as data from social networks such as Facebook and Twitter.[88] The unobservable cases of interest are the location and time of future crimes, the likely perpetrators or victims of future crimes, and likely suspects in past crimes.

One report "found a near one-to-one correspondence between conventional crime analysis and investigative methods and the more recent 'predictive analytics' methods"[89]—which is to say that, for the most part, police methods have not changed, but the predictive analytics substitute for the older modes of analysis. While police seek the same sorts of answers as before, data mining allows them to find patterns that they could not have otherwise discovered on their own.[90] Police have long understood that some crime is hyperlocal, resulting in the formation of hotspots.[91] Offender profiling has its basis in the psychology of crime and criminals, although its efficacy has never truly been clear.[92] But whereas past forms of prediction relied on some theory of

---

[87] PERRY ET AL., *supra* note 9, at 2; *see generally* Chris Jay Hoofnagle, *Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement*, 29 N.C. J. INT'L L. & COM. REG. 595 (2004).

[88] *See* John Buntin, *Social Media Transforms the Way Chicago Fights Gang Violence*, GOVERNING (Oct. 2013), http://www.governing.com/topics/urban/gov-social-media-transforms-chicago-policing.html (describing the use of social media data in policing techniques in Chicago); Johnson et al., *supra* note 8 (explaining that the use of social network analysis is "a valuable tool for law enforcement"); Craig Timberg & Elizabeth Dwoskin, *Facebook, Twitter And Instagram Sent Feeds That Helped Police Track Minorities In Ferguson And Baltimore, Report Says*, WASH. POST (Oct. 11, 2016), https://www.washingtonpost.com/news/the-switch/wp/2016/10/11/facebook-twitter-and-instagram-sent-feeds-that-helped-police-track-minorities-in-ferguson-and-baltimore-aclu-says/.

[89] PERRY ET AL., *supra* note 9, at xiv.

[90] *See* Bachner, *supra* note 8, at 17 ("[O]ne of the key benefits of predictive policing is that previously unknown or overlooked patterns [in the raw data] emerge . . . .").

[91] *See* Ferguson, *supra* note 7, at 273–76 (discussing criminology theories of "repeated patterns of localized crime").

[92] *See* AINSWORTH, *supra* note 78, at 176–78 (noting that the accuracy or profiling is largely unknown and that "even if it were possible" to determine, "accuracy does not equate with utility").

criminology, data mining allows the same correlative principles to be expanded more broadly. Data mining allows police to operate unconstrained by theory, finding correlations without worrying about why they work.[93]

The majority of predictive policing systems in use are either "place-based" systems that aim to predict when and where future crime will occur, or "person-based" systems that attempt to predict offenders, determine the identities of perpetrators, or predict potential victims.[94] Person-based systems can be separated further into those used to solve a particular crime—what I call "suspect-based" systems[95]—or those used to assess individuals' threat levels in the abstract.[96] Some technical detail about how the systems work will be important to understanding why, without care, they are likely to result in disparate impacts against vulnerable communities. That discussion follows.

1. Place-Based Predictive Policing. Place-based systems, including the well-known examples of software from PredPol and HunchLab, are the most common type of predictive policing.[97] While this discussion begins with place-based predictive policing, the three types are not so fundamentally different. Understanding the pitfalls of one is key to understanding where all three types of predictive policing can go wrong. Ultimately, however, the

---

[93] *See* VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA 55 (2013) ("No longer do we necessarily require a valid substantive hypothesis about a phenomenon to begin to understand our world."); Rob Kitchin, *Big Data, New Epistemologies and Paradigm Shifts*, BIG DATA & SOC'Y, Apr.–June 2014, at 1, 2 ("Big Data analytics enables an entirely new epistemological approach for making sense of the world; rather than testing a theory by analysing relevant data, new data analytics seek to gain insights 'born from the data.' "). A complete absence of theory from data mining is actually a bit of an exaggeration. Background assumptions are required simply for machine learning to work, but those assumptions need not be detailed. *See* Domingos, *supra* note 85, at 81 (explaining that "[e]very learner must embody some knowledge or assumptions beyond the data it is given in order to generalize beyond it," but that "very general assumptions . . . are often enough to do very well").

[94] *See* PERRY ET AL., *supra* note 9, at 8 (explaining that some approaches are used "to forecast places and times with an increased risk of crime" and others to "identify individuals at risk of offending in the future"); ROBINSON & KOEPKE, *supra* note 18, at 2; *see also* Ferguson, *supra* note 12, at 1126–43 (discussing the iterations of predictive policing "1.0" (place-based property crime), "2.0" (place-based violent crime), and "3.0" (person-based crime)).

[95] This is my term, designed to distinguish from "person-based."

[96] *See* ROBINSON & KOEPKE, *supra* note 18, at 3 (describing person-based systems as those "predicting the identities of people particularly likely to commit . . . certain kinds of crime").

[97] *See id.* at 3–4.

disparate impact harm of each is different in both form and degree.

Place-based predictions are primarily focused on hot-spot detection, which, in turn, is used mostly for resource management.[98] Police want to put more officers where crime is occurring. Occasionally, if there is a very specific pattern, the police may be able to predict the next instance in a crime spree, but the tools are not usually that specific.

The potential for harm stemming from racially imbalanced outcomes is the harm resulting from having more police in a neighborhood that is unfairly maligned as having more crime. If one believes that all crime should lead to arrest, one may not readily see the harm from over-policing nonwhite neighborhoods.[99] After all, if the restrictions of the Fourth Amendment are observed, only people likely committing crimes will be arrested. But that picture is not the reality. First, a great deal of crimes—such as minor drug use or public intoxication—do not always lead to arrest. They are often not observed, or if observed, an arrest is subject to police discretion.[100] Unsurprisingly, these arrests for such crimes are more common among people of color.[101] Second, even for non-nuisance crimes, police have limited resources. Hypothetically, if a city has two racially segregated neighborhoods—one black, one white—and has enough of a murder problem that not all of them in either neighborhood can be solved, a police policy that focuses entirely on the black neighborhood would not be normatively acceptable, even if all the people the police arrest individually deserve to be arrested. Thus, on a systemic level, over-policing nonwhite neighborhoods does present a fairness harm.

So how does it happen? There are several mechanisms by which this type of data mining system could result in a disparate impact

---

[98] *E.g.*, *About PredPol*, PREDPOL, http://www.predpol.com/about/ (last visited Oct. 25, 2017) ("PredPol aims to keep communities safer. Our day-to-day operations tool identifies where and when crime is most likely to occur so that you can effectively allocate your patrol resources and prevent crime.").

[99] For an excellent account of the devastation over-policing and mass incarceration have caused communities of color, see MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS (2012).

[100] *See* ROBINSON & KOEPKE, *supra* note 18, at 5 (citing the disparate arrest rates for white and black marijuana users that is caused "in part . . . [because] police exercise an extraordinary degree of discretion in deciding what to report as crimes").

[101] *See* ALEXANDER, *supra* note 99, at 98–99, 185.

on protected classes.[102] These mechanisms correspond to the different steps in the workflow:[103] (1) designing the problem; (2) collecting the training data and labeling examples within it; (3) selecting features to model; and (4) the potential for accidentally using proxies for a protected class.[104] Each of these mechanisms has an application to the realm of predictive policing.

First, data miners must define the problem in a way that a computer can understand.[105] An officer cannot merely ask a computer, "How can I prevent crime?" Rather, the officer must take an amorphous question about the world and translate it such that the outcome can be expressed as "the value of some target variable."[106] For example, PredPol divides a map into 500ft ×

---

[102] *See* Barocas & Selbst, *supra* note 19, at 677–93. Of the various versions of predictive policing, a suspect-based system is the most related to those used in the employment context in which we wrote. There, an employer will have a model of what a person with good or bad outcomes looks like and will test that model against a particular candidate. The same is true when the police look for a suspect—thus, the concerns are the same.

[103] This is actually the workflow of a supervised learning system. For the purposes of this Article, I will only discuss supervised learning because that is what is overwhelmingly used today. *See* COMM. ON THE ANALYSIS OF MASSIVE DATA ET AL., NAT'L RESEARCH COUNCIL, FRONTIERS IN MASSIVE DATA ANALYSIS 104 (2013), http://www.nap.edu/catalog.php?record_id=18374 (noting that "[p]redictive modeling is referred to as supervised learning in the machine-learning literature"). Supervised learning techniques include classification, estimation, and prediction. *Id.* at 104–06. Forms of data mining that involve sorting or ranking of outcomes involve supervised learning. *Id.* at 115. Unsupervised learning includes techniques such as clustering, which means grouping elements of a set based on similarity without specifying any particular outcome beforehand. *Id.* at 102. Thus, if all crime in a city over a period of time were plotted, and a data miner specified a certain number of clusters, the algorithm would determine where the crimes were most tightly focused. *Id.* at 103. The primary difference between supervised and unsupervised learning is whether the data miner seeks a value of a target variable, or instead wants to find something interesting about the data without specifying ahead of time what is sought. *Id.* at 101. While unsupervised learning techniques suffer from many of the same pitfalls as supervised learning, the difficulties with problem definition look different because unsupervised learning does not solve a specified problem in the same way. There is also a third type of machine learning, known as "reinforcement learning," in which machines are able to interact with the world separately from human involvement and learn from their interventions. However, it is not used in predictive policing systems. *See* M.I. Jordan & T.M. Mitchell, *Machine Learning: Trends, Perspectives, and Prospects*, SCIENCE, July 17, 2015, at 258, http://science.sciencemag.org/content/sci/349/6245/255.full.pdf.

[104] *See* Barocas & Selbst, *supra* note 19, at 677–93; *see also* PETE CHAPMAN ET AL., CRISP-DM 1.0: STEP-BY-STEP DATA MINING GUIDE 10–12 (2000).

[105] Barocas & Selbst, *supra* note 19, at 678.

[106] *Id.*

500ft squares, and for each, the target variable becomes the likelihood of a given crime.[107]

But the categories are not always obvious. If the system is designed to detect crimes within a particular square on a map, it should separate out types of crime. Should "type of crime" be broken down into violent and nonviolent? Should property crimes or nuisance crimes be counted separately? Should nuisance crimes then be further broken down? Deciding how to parse the problem can have severe consequences for the ultimate outcome.[108] For example, if the nuance between robberies and burglaries is missing because both are placed in the "property crime" bucket, the algorithm may not detect the difference between an area with high amounts of robberies and an area with a high number of burglaries, though the two crimes might be perpetrated by different people with different victims.

Using data mining also tends to bias organizations toward questions that are easier for computers to understand.[109] Property crime prediction is a common goal of predictive policing because "[b]urglars tend to be territorial,"[110] and geographic analyses are relatively easy to create. A police department is thus likely to focus more on property crime than it otherwise would.[111] It is doubtful that the demographics of property crime are exactly the same as other crimes, so skewing the system in this way not only affects the absolute ratio of type of crimes policed, but also the demographics. A similar bias may result from what systems are commercially available. If companies offer systems designed to

---

[107] *Technology*, PREDPOL, http://www.predpol.com/technology/ (last visited Sept. 19, 2017) (describing in detail the technology used by PredPol).

[108] *See* Oscar H. Gandy Jr., *Engaging Rational Discrimination: Exploring Reasons for Placing Regulatory Constraints on Decision Support Systems*, 12 ETHICS & INFO. TECH. 29, 38 (2010).

[109] *See, e.g.*, Jon Kleinberg et al., *Prediction Policy Problems*, 105 AM. ECON. REV. 491, 494 ("[I]mproved prediction using machine learning techniques can have large policy impacts . . . [but e]ven this small set of examples are biased by what we *imagine* to be predictable.").

[110] Nate Berg, *Predicting Crime, LAPD-Style*, THE GUARDIAN (June 25, 2014), https://www.theguardian.com/cities/2014/jun/25/predicting-crime-lapd-los-angeles-police-data-analysis-algorithm-minority-report (quoting Captain John Romero of the LAPD Real-Time Analysis and Critical Response Division).

[111] *See* Elizabeth E. Joh, *Policing By Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 58 (2014). The same is likely true of nuisance crimes. CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 89 (2016).

detect property crimes, but other crimes have no similar system, law enforcement are likely to choose to police the former. In early 2017, this bias led Sam Lavigne, Francis Tseng, and Brian Clifton to create a satirical, white-collar crime predictor based on the same place-based predictive policing methods that could be used for street crime.[112]

The next biases come from the training data. A data mining system learns by example, and must take its training data as "ground truth," because that data is the only information the algorithm has about the outside world.[113] A big part of getting the data right is correctly labeling the examples the algorithm is trained on.[114] The most common source of data for predictive policing algorithms—used in every version of predictive policing in existence—is past crime data, often collected by the police themselves.[115] Therefore, in predictive policing, labeling examples will commonly mean determining whether a data point was or was not a crime, and if so, what type of crime it was. Reliance on past data is problematic, though, as accurate crime data rarely exists.[116]

There are several reasons for this, but one major reason is that the most systematic contact police departments have with "criminals" is at the moment of arrest.[117] Results after arrest are

---

[112] Sam Lavigne, Francis Tseng & Brian Clifton, *White Collar Crime Risk Zones*, THE NEW INQUIRY (Apr. 26, 2017), https://thenewinquiry.com/white-collar-crime-risk-zones/. Predictive systems also exist for financial crimes, but those crimes might be under the jurisdiction of another agency such as the Securities and Exchange Commission. *See* Mary Jo White, Chair, SEC, Keynote Address at the 41st Annual Securities Regulation Institute (Jan. 27, 2014) (transcript available at http://www.sec.gov/News/Speech/Detail/Speech/1370540677500) (describing the SEC's NEAT program as using data mining to identify insider trading).

[113] Barocas & Selbst, *supra* note 19, at 682.

[114] *Id.*

[115] ROBINSON & KOEPKE, *supra* note 18, at 3–4 (listing predictive policing systems and noting that historical crime data is used in all of them).

[116] *See* DELBERT S. ELLIOT, CTR. FOR THE STUDY AND PREVENTION OF VIOLENCE, LIES, DAMN LIES AND ARREST STATISTICS 1 (1995); DAVID A. HARRIS, PROFILES IN INJUSTICE: WHY RACIAL PROFILING CANNOT WORK 75–78 (2002); Lawrence W. Sherman & Barry D. Glick, *The Quality of Police Arrest Statistics*, POLICE FOUND. REP., Aug. 1984, at 1.

[117] *See* HARRIS, *supra* note 116, at 77 (explaining that arrest rates are poor measures of criminal activity and are instead merely measures of law enforcement activity, as arrest rates only indicate contact with police and do not fully and accurately depict all offenders or instances of criminal activity). Reporting of crime and suspicious activity by citizens is also biased. *See* Jason Tashea, *Websites and Apps for Sharing Crime and Safety Data Have*

often not updated.[118] Thus, most research in crime statistics[119] and most actuarially-driven criminal justice systems[120] use arrest data as the best available proxy, even though arrests are racially biased[121] and in other ways a poor proxy for crime.[122] Even if post-arrest statistics were collected, a great number of cases end in plea agreements that do not reflect the crime the arrestee committed or was originally arrested for—thus, using these statistics would not solve the problem either.[123] As a result, a majority of crime labels may be incorrect, whether describing a type of crime or the existence of one, and thus models will learn that people of color commit a higher percentage of "crimes" than they do in reality. It is worth emphasizing again that—due to "redundant encodings" in the data sets—a model need not have race as an input to correlate an output with race. It will simply detect those other encodings that are good real-world proxies for race and rely on them instead.[124]

Training data must also be a representative sample of the whole population. The ultimate goal of data mining is pattern-matching and generalization,[125] and without a representative sample, generalizing introduces sampling bias.[126] There are many potential sources of sampling bias. For example, data can be skewed by past historical practices. When police allocate more resources in areas where there has been more crime in the past,

---

Become Outlets for Racial Profiling, A.B.A. J. (Aug. 1, 2016), http://www.abajournal.com/magazine/article/crime_safety_website_racial_profiling.

[118] ELLIOT, *supra* note 116, at 2.

[119] *Id.* at 1.

[120] *See* Jessica M. Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59, 98 (2017); *see also* Eisha Jain, *Arrests as Regulation*, 67 STAN. L. REV. 809, 815 (2015) (describing how arrests are used as proxies outside the criminal context "not necessarily because they are the best screening tools," but "because they are relatively easy and inexpensive to access and because [noncriminal justice actors] regard arrests as proxies for information they value, such as the potential for violence, unreliability, or instability").

[121] *See, e.g.*, Tracey Maclin, *The Fourth Amendment on the Freeway*, 3 RUTGERS RACE & L. REV. 117, 121 (2001).

[122] Jain, *supra* note 120, at 832.

[123] Eaglin, *supra* note 120, at 101–04.

[124] *See* Hardt, *supra* note 23 (describing how machine learning often results in the program "predicting unknown protected attributes for other seemingly innocuous features").

[125] Domingos, *supra* note 85, at 80.

[126] *See* Barocas & Selbst, *supra* note 19, at 686.

crimes in those areas will be over-represented in future data.[127] This error is the best-understood version of algorithmic discrimination as applied in the policing context. Bernard Harcourt discussed it in his 2007 book arguing against actuarial policing.[128] Worse yet, when predictive policing is used specifically to figure out where to put more officers, this phenomenon creates a positive feedback loop that further skews future data, as the increased police presence will lead to detection of more crimes in that area.[129] Reporters and advocates have recognized these dangers.[130] An August 2016 statement released by seventeen civil rights organizations noted that predictive policing would be inherently biased because of its reliance on past crime data that "primarily document[s] law enforcement's *response* to the reports they receive and situations they encounter, rather than providing a consistent or complete record of all the crimes that occur."[131]

Another source of discriminatory effect is feature selection. Data miners must "make choices about what attributes they observe and subsequently fold into their analyses."[132] By necessity, the police

---

[127] Scott L. Johnson, *The Self-Fulfilling Prophecy of Police Profiles*, *in* THE SYSTEM IN BLACK AND WHITE: EXPLORING THE CONNECTIONS BETWEEN RACE, CRIME, AND JUSTICE 93, 105 (Michael W. Markowitz & Delores D. Jones-Brown eds., 2000) (noting that "profiles legitimate past bias . . . which then generates more bias"); William J. Stuntz, *The Political Constitution of Criminal Justice*, 119 HARV. L. REV. 781, 791 (2006) ("Whom the police catch depends on where they look.").

[128] *See* HARCOURT, *supra* note 52, at 147–50 ("[I]f the police dedicate more resources to investigating, searching, and arresting members of a higher-offending group, the resulting distribution of arrests . . . will disproportionately represent members of that higher-offending group.").

[129] Ensign et al., *supra* note 47; *see also* HARCOURT, *supra* note 52, at 147–50 (discussing what he calls the "ratchet effect").

[130] *See* STATEMENT OF CIVIL RIGHTS GROUPS, *supra* note 20 (stating the position of various advocacy groups that predictive policing will only exacerbate the disproportionate scrutiny that minority communities are subjected to from law enforcement); Bryan Llenas, *Brave New World of 'Predictive Policing' Raises Specter of High-Tech Racial Profiling*, FOX NEWS (Feb. 25, 2014), http://www.foxnews.com/world/2014/02/24/brave-new-world-predictive-policing-raises-specter-high-tech-racial-profiling.html (quoting attorney and activist Hanni Fakhoury as saying "if the data is biased to begin with and based on human judgment, then the results the algorithm is going to spit out will reflect those biases"); Matt Stroud, *The Minority Report: Chicago's New Police Computer Predicts Crimes, But Is It Racist?*, THE VERGE (Feb. 19, 2014), https://www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist (citing "red flags" that have been raised by Chicago's predictive policing system).

[131] STATEMENT OF CIVIL RIGHTS GROUPS, *supra* note 20.

[132] Barocas & Selbst, *supra* note 19, at 688.

make judgment calls about where geographic hot spots are, what features they should aim to contain, and how big they should be.[133] They must also decide whether to remain simple and take into account only location, crime type, and date and time, as PredPol does,[134] or include many other variables like socioeconomic indicators, weather, seasonality, recurring events and holidays, and proximity of other known offenders—as is the case with a product like HunchLab.[135] These choices have downstream effects.

The possibility of error rates is exacerbated if police or the software companies they contract with add features by purchasing data from data brokers. Their profiles often are not correct[136] and, at best, are optimized for commercial uses, not police work.[137] Data brokers assemble these profiles with the assumption that they will be used for targeted advertising,[138] where the total stakes for an errant profile is the risk that someone sees an incorrect advertisement. Data brokers' incentives are to make their models just good enough so that their customers can profit more by using them than by not using them.[139] This is a very error-tolerant metric. There is no reason to suspect that low absolute error rates are even of interest to commercial data brokers—that they will not, for example, link information to the wrong person,[140] or that they have any interest in assuring the representativeness of their data sets.[141]

---

[133] *See* Bachner, *supra* note 8, at 20.

[134] *About PredPol*, PREDPOL, http://www.predpol.com/about/ (last visited Sept. 19, 2017).

[135] AZAVEA, HUNCHLAB: UNDER THE HOOD 12 (2015), https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf.

[136] *See* Bobby Allyn, *How the Careless Errors of Credit Reporting Agencies Are Ruining People's Lives*, WASH. POST (Sept. 8, 2016), https://www.washingtonpost.com/posteverything/wp/2016/09/08/how-the-careless-errors-of-credit-reporting-agencies-are-ruining-peoples-lives/.

[137] FED. TRADE COMM'N, DATA BROKERS: A CALL FOR TRANSPARENCY AND ACCOUNTABILITY 38 (2014), https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf.

[138] *See id.* at 26–27.

[139] *See id.* at 36 ("The procedures that the data brokers use to assure the quality of the data they provide to clients depend on the type of product at issue and the data broker's business model.").

[140] *See* Jain, *supra* note 120, at 832 ("Arrest data may be linked to the wrong person—particularly when arrested individuals have common names or provide false identification at the time of their arrest.").

[141] *See* O'NEIL, *supra* note 111, at 12–13.

Finally, the coarseness or granularity of features could affect outcomes along the lines of protected class. Features at the wrong level of granularity can result in generalizations that are "simultaneously rational and unfair" because certain individuals are "actuarially saddled" by statistically sound inferences that are nevertheless inaccurate.[142] That is, proximity or similarity to certain groups of outcomes will cause an inappropriate adverse determination. But feature selection is unavoidably subjective, and it is often unclear beforehand whether is it more accurate or unfair to define the location of a crime by address, 500-foot square, city block, or square mile. Moreover, the more "accurate" decision may not lead to the fairest result for people swept up in that region.

2. Person-Based Predictive Policing. The next type of predictive policing is person-based, but not investigation driven. For example, Intrado's Beware software allows police to draw on publicly available data, including social media data, to check the "threat score" of a person or address as a 911 call comes in, and to assign a label of green, yellow, or red, accordingly.[143] Other systems analyze social media to automatically find gang members.[144] Still other systems, like Chicago's "heat list," find the likeliest people to be involved in an unspecified future crime.[145]

The disparate impact harm stemming from racial imbalance in these systems is different. These systems could lead to extra monitoring of their subjects, and when a later crime occurs, police might be more likely to look at them first. Or, if police respond to a call with an erroneous "red" threat level, they might proceed anxiously—with an itchy trigger finger—or otherwise be more easily provoked into unnecessary force. Because the effects of these systems are aimed at individuals,[146] the harm also looks

---

[142] FREDERICK SCHAUER, PROFILES, PROBABILITIES, AND STEREOTYPES 3–7 (2006).

[143] *See* Justin Jouvenal, *The New Way Police Are Surveilling You: Calculating Your Threat 'Score,'* WASH. POST (Jan. 10, 2016), https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c_story.html.

[144] *See, e.g.*, Lakshika Balasuriya et al., *Finding Street Gang Members on Twitter* 2016 IEEE/ACM INT'L CONF. ON ADVANCES IN SOC. NETWORKS ANALYSIS & MINING (ASONAM), 685, https://arxiv.org/pdf/1610.09516v1.pdf; Buntin, *supra* note 88.

[145] *See* Buntin, *supra* note 88.

[146] *See id.*

different than the results of resource-management decisions driven by crime-mapping. But some of the effects can be similar in scale. As sociologist Sarah Brayne has documented, the Los Angeles Police Department's person-based predictive policing uses a simple points-based system, where more points means a person is a greater threat.[147] To find "the worst of the worst," the LAPD adds one point per police contact, leading to the very same type of feedback loops that exist in place-based policing.[148] Over time, the erroneous appearance of greater threat levels in minority neighborhoods could also exacerbate an already adversarial relationship with police and endanger lives as a result.

Just like mislabeled instances of crime in place-based systems, the embedding of historically biased policing will teach the algorithm that being a person of color makes one more likely to be a criminal.[149] Beyond historical inaccuracies, person-based systems are likely to encounter more data collection pitfalls than place-based systems. For example, when training an algorithm with examples of people who have been shot and have not been shot, the police have much more data on those who have been shot. While this "class imbalance" is a fixable problem in principle,[150] care must be taken to ensure the representativeness of a much larger class of people who have not been shot, so as to avoid sampling bias. Social media data is also vulnerable. The structural biases of the particular system the police extract data from, whether Twitter, Facebook, or some other service, could change the patterns of connections that are observed.[151] Whether police attempt to extract a generalizable pattern of associations[152]

---

[147] *See* Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 82 AM. SOC. REV. 977, 986–89 (2017).

[148] *Id.* at 987 ("An individual having a high point value is predictive of future police contact, and that police contact further increases the individual's point value.").

[149] *See id.* at 997.

[150] *See* Jason Brownlee, *8 Tactics to Combat Imbalanced Classes in Your Machine Learning Dataset*, MACHINE LEARNING MASTERY (Aug. 19, 2015), http://machinelearningmastery.com/tactics-to-combat-imbalanced-classes-in-your-machine-learning-dataset/.

[151] Zeynep Tufekci, *Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls*, PROC. EIGHTH INT'L AAAI CONF. WEBLOGS & SOC. MEDIA, 2014, at 505, 508, http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8062/8151.

[152] *See* Bachner, *supra* note 8, at 22–24.

or perform a social media analysis,[153] such as with an attempt to find gang members,[154] they have to understand the ways that the social networking platform changes the data from what they might expect to see in the offline world.

Feature selection is also more complicated with respect to people rather than places. Representations of people in data are necessarily reductive. As Toon Calders and Indrė Žliobaitė have noted, "[i]t is often impossible to collect all the attributes of a subject or take all the environmental factors into account with a model."[155] Police may be tempted to use certain types of data—for example, race, gender, neighborhood, or age—because it is easily accessible. Choice of features would ideally not be made based on cost or accessibility. Features that do not adequately capture the relevant distinctions between people or locations will make the predictions less accurate. But cost and convenience are common factors in these decisions, and both can lead to discriminatory outcomes.[156]

3. Suspect-Based Predictive Policing. The final type of system is suspect-based. Suspect-based systems are the digital descendants of offender profiling. They will be used to create a model for what a person who might commit a particular crime might look like, and then that model will be used to locate suspects.[157] Though not yet commonly deployed—at least as far as one can tell from public information—suspect-based predictive

---

[153] *See* Somini Sengupta, *In Hot Pursuit of Numbers to Ward Off Crime*, N.Y. TIMES (June 19, 2013), http://bits.blogs.nytimes.com/2013/06/19/in-hot-pursuit-of-numbers-to-ward-off-crime/.

[154] *See* Emerging Technology from the arXiv, *How to Detect Criminal Gangs Using Mobile Phone Data*, MIT TECH. REV. (Apr. 14, 2014), https://www.technologyreview.com/s/526471/how-to-detect-criminal-gangs-using-mobile-phone-data/ (describing a software platform "that can bring together information . . . to recreate detailed networks behind criminal organizations").

[155] Toon Calders & Indrė Žliobaitė, *Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY 43, 47 (Bart Custers et al. eds., 2013).

[156] *See id.* at 52 (noting that data features are often not collected because the "data is hard to collect," which results in "overestimating the importance" of the features that are collected).

[157] *See* Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 876 (2016) (describing "Automated Suspicion Algorithms" that "seek to predict individual criminality").

policing systems will be here soon.[158] They are also the most troubling of the three types.

Here, the harm is that racial disparities in the outcome of the algorithm create a greater degree of suspicion and higher likelihood of finding probable cause due to a suspect's race. While the mechanisms of discrimination are similar to those above, it is worth separating out suspect-based policing because it has most vividly captured the imagination of those writing about predictive policing,[159] and it is the most likely to respond to Fourth Amendment oversight. That is because, unlike the prior two methods, suspect-based policing would be used in service of an investigation, which is the primary context in which the Fourth Amendment operates.[160] Andrew Ferguson has pointed out troubling difficulties with using the Fourth Amendment to address Big Data-driven investigations,[161] but at least it is not a total conceptual mismatch.

C. BUT IS IT ALWAYS DISCRIMINATION?

While there are many ways an algorithm could be skewed in a direction harmful to protected classes, the algorithm could also be accurate and still have a disproportionate impact. The model could have no data quality problems and "optimal" choices of problem definitions and features, but still make determinations primarily based on a trait or group of traits that, due to redundant encodings, incidentally serves as a proxy for race. Here, the algorithm would be rediscovering certain inequalities in society

---

[158] *See* Ferguson, *supra* note 25, at 351; Rich, *supra* note 157, at 871–75, 878–79; *see also* Reed E. Hundt, *Making No Secrets About It*, 10 ISJLP 581, 588 (2014) ("[G]overnment now routinely asks computers to suggest who has committed crimes.").

[159] *See, e.g.*, Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15, 28 (2016) (discussing "algorithms that search for suspicious activity"); Rich, *supra* note 157, at 897 (discussing "automated suspicion algorithms").

[160] *See* Renan, *supra* note 33, at 1042–43 (explaining that the Fourth Amendment does not address programmatic surveillance because it is transactional—that is, focused on individual moments in investigations); Meares, *supra* note 32, at 165 (noting that "the constitutional framework is based on a one-off investigative incident").

[161] *See* Ferguson, *supra* note 27, at 401–04 (explaining that using big data as a justification to stop individuals will "undermin[e] the individualized and particularized protections in the Fourth Amendment").

that lead to disparate rates of crime among people of different groups.

In this case, the very accuracy of the determination would cause the racial disparity in output.[162] This leads to some difficult questions. If there is a neighborhood or hot spot that truly does have more crime, is it actually discriminatory to have more police presence there? Is it even harmful? The communities in which more crimes occur might welcome a greater police presence. Even if they do not welcome it, though, it may not be "discriminatory," as we usually use the term. On the one hand, if police increase their presence in communities that are already in bad shape, that may increase the cycle of community disruption and poverty, and exacerbate the extant criminal problem.[163] On the other hand, the police are just doing their job. Even if there is agreement that such a result is unfair, fixing it would require the police to make less accurate determinations in order to racially rebalance the algorithm.[164] Asking police to catch fewer criminals after conceding the accuracy of their algorithms would be a hard sell.

Nonetheless, there are reasons to be cautious about this conclusion. First, it will often be difficult to observe a disparate impact in the output of such a system and determine conclusively whether it is actually a reflection of reality, or a function of the various problems described above.[165] Without some form of perfectly omniscient data, this may be functionally impossible.[166]

---

[162] *See* Barocas & Selbst, *supra* note 19, at 692.

[163] *See* Ferguson, *supra* note 76, at 230 ("The counterintuitive result is that a greater police presence can, in fact, foster the social conditions that increase crime. Disrupting existing social connections through arrest, incarceration, or intrusive surveillance causes normal social connections break down.").

[164] *See* Barocas & Selbst, *supra* note 19, at 692 (noting that, because a "more precise form of data mining will be more likely to capture disparate impact, police would have to utilize less accurate data to resolve the problem").

[165] *See infra* Part IV.A.

[166] A group of computer scientists, recognizing that problems where good "ground truth" data is available are fundamentally different from those where it is not, have proposed different technical fairness measures for those situations. Where ground truth data is available, they propose a measure called "disparate mistreatment" that aims to equalize error rates of the prediction between groups. Muhammad Bilal Zafar et al., *Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification Without Disparate Mistreatment*, Proc. 26th Int'l World Wide Web Conf. 1171, 1171 (2017). This is a better measure because, once equalized, error rates are made an optimization constraint, and the improvements in the algorithms will benefit everyone equally. But where ground truth data does not exist or is untrustworthy, they advocate a return to the disparate impact

Scholars and advocates have made this argument repeatedly. They argue that, although police officers see arrest rates as gospel about differing demographic crime rates, there is no such proof.[167] And as discussed in greater depth later, the realities of data mining often make it impossible to tell whether the ultimate source of discrimination is error or reality.[168] Thus, because the claim that discrimination merely reflects reality may not be sustainable, the "hard sell" might not turn out to be a realistic scenario.

Second, even if there is no discrimination in the legal sense, there is still a broader fairness argument to be made against the result. If the algorithms are mere reflections of reality, then another route to addressing crime is addressing the background conditions of these communities that lead to increased crime.[169] Police need not use these systems solely for criminal enforcement. For example, when the Chicago police rolled out plans for the Strategic Subjects List, they claimed that it would lead at least partly to the provision of social services.[170] This makes sense, as the list predicts both victims and perpetrators of gun crime.[171] If the results of predictive policing are used to help those deemed likely to be involved in future crime, as a preventative measure, then even skewed data should not be thought of as discriminatory because there is no harmful result. But the current evidence demonstrates that rather than involving social services, "the prevention strategy . . . was not well developed and only led to increased contact with a group of people already in relatively

---

standard, *id.* at 1172, which in computer science means aiming equalize *outcomes* rather than error rates within an acceptable margin of disparity.

[167] *See, e.g.*, HARRIS, *supra* note 116, at 75–78.

[168] *See* Barocas & Selbst, *supra* note 19, at 722; *infra* Part IV.A.

[169] *See* ROBINSON & KOEPKE, *supra* note 18, at 6 (suggesting that data could be used to "track and reward strategies that do a better job of balancing a community's needs and interests").

[170] *Id.* at 9.

[171] *Id.* ("According to one newspaper report, this was meant to be a carrot-and-stick approach, where individuals on the list would be warned that 'further criminal activity, even for the most petty offenses, will result in the full force of the law being brought down on them . . . At the same time, police extend them an olive branch of sorts, an offer of help obtaining a job or of social services.' " (quoting Jeremy Gorner, *Chicago Police Use 'Heat List' as Strategy to Prevent Violence*, CHI. TRIB. (Aug. 21, 2013), http://articles.chicagotribune.com/2013-08-21/news/ct-met-heat-list-20130821_1_chicago-police-commander-andrew-papachristos-heat-list)).

frequent contact with police."[172] A statement released by a Consortium of Civil Rights groups stated, "[o]ther vital goals of policing, such as building community trust, eliminating the use of excessive force, and reducing other coercive tactics, are currently not measured and not accounted for by these systems."[173] This accords with Kevin Haggerty and Richard Ericson's observations at the turn of the century that even efforts by social service agencies "wedded to a welfarist ideology of service delivery" were later "drawn into the harder edge of social control."[174] As Sarah Brayne similarly observed, "regardless of the reason they were kept in the first place, data and records are increasingly integrated and deployed by law enforcement agencies for a broad range of surveillance purposes."[175]

If police see their job as surveilling and arresting criminals, and see predictive policing as merely a way to identify them or determine places and times at which they can arrest them, then these systems will produce the unfair results described above,[176] even when accurate. But this need not be the way predictive policing is used. Rachel Harmon has argued that police usually define their job by arrests, but that such an extreme focus on arrests has unexamined and unjustified costs.[177] Arguing that police should have less discretion to arrest, she writes that "[i]f more people can, through a less discretionary process, be released with only a low increase in failures to appear and reoffending, then broad discretion to arrest is no longer justified."[178] Though she does not discuss predictive policing, this idea of detecting risk is the central purpose of predictive policing technology. If the technology is genuinely demonstrating facts about current reality, then aiming to change that reality rather than perpetuating it through arrest is the fairer result.

---

[172] Saunders et al., *supra* note 18, at 363.
[173] STATEMENT OF CIVIL RIGHTS GROUPS, *supra* note 20.
[174] Kevin D. Haggerty & Richard V. Ericson, *The Surveillant Assemblage*, 51 BRIT. J. SOC. 605, 611 (2000).
[175] Sarah Brayne, *Surveillance and System Avoidance: Criminal Justice Contact and Institutional Attachment*, 79 AM. SOC. REV. 367, 371 (2014).
[176] *See supra* notes 172–82 and accompanying text.
[177] *See* Rachel A. Harmon, *Why Arrest?*, 115 MICH. L. REV. 307, 313–20 (2016).
[178] *Id.* at 354.

III. THE FAILURES OF STANDARD ACCOUNTABILITY TOOLS

While data mining has the ability to be superior to police hunches and reduce discriminatory results, police cannot simply trust that their analytics are accomplishing those tasks. And just as police cannot trust their tools not to be discriminatory, society cannot merely trust police to know or care. The adoption of new, potentially harmful policing tools must be regulated somehow. But our standard modes of regulation are not working. Typically, police are regulated through the Fourth and Fourteenth Amendments, but neither has much to say about unintentional discrimination. Even if they did, or a hypothetical anti-discrimination law was passed, it would pose its own challenges. Thus, before Part IV proposes a new model of regulation to address the potential for discrimination in predictive policing, this Part explains why the current models fail.

A. THE CONSTITUTION DOES NOT PROVIDE THE ANSWER

If the police arrest a person on the recommendation of a suspect-based predictive policing algorithm, the arrestee might hope that the Fourth Amendment can provide a solution. The arrest would go something like this: police are driving down the street, running a facial recognition program to identify people,[179] and then running those names through their algorithms based on publicly available data to see who matches a profile. Once they find a match, they arrest the person on suspicion of whatever crime they are looking to solve. This scenario is a quintessential Fourth Amendment problem.

Andrew Ferguson has argued that such an arrest would not raise Fourth Amendment concerns on its own.[180] Yet the Fourth Amendment is the primary tool for police regulation in American

---

[179] *See* Clare Garvie et al., *The Perpetual Line-Up: Unregulated Police Face Recognition in America*, GEORGETOWN CTR. ON PRIVACY & TECH. (Oct. 18, 2016), https://www.perpetualline up.org.

[180] *See* Ferguson, *supra* note 25, at 330 (describing a hypothetical scenario using facial recognition software in which the police attain "particularized, individualized suspicion about a man who is not doing anything overtly criminal").

law.[181] Thus, it is worth pushing a little further to ask a slightly different question: Could the Fourth Amendment be implicated if the algorithm was shown to have racially disparate results? An arrestee in the above scenario might argue that he was unreasonably searched or seized because the police methods were discriminatory or that the reliance on race implies a lack of individualized suspicion. Person-based tools have a more tenuous connection to the Fourth Amendment because even if a person is watched more closely, that should, in theory, be separate from the facts leading to probable cause. And place-based tools used for resource management will not create a Fourth Amendment concern, because those tools are not related to investigations.

But it turns out that the Fourth Amendment will not address the potential harms identified in Part II. Nor will the Equal Protection Clause of the Fourteenth Amendment, because it does not protect against disparate impact.[182] The standard strategies for constitutional regulation of the police are therefore ill-suited to address disparate impact caused by predictive policing.

1. Race and the Fourth Amendment. Discussions of race and the Fourth Amendment usually begin with *Whren v. United States*,[183] a 1996 case holding that the subjective motives of police officers, including racial bias, do not invalidate an otherwise lawful stop.[184] In *Whren*, a police officer stopped two black men in an SUV in a "high drug area" of Washington D.C., and found drugs in the car.[185] Moving to suppress the evidence, the defendants argued that a reasonable police officer would not have stopped them for the stated reasons, and that those reasons were mere pretext for a racially motivated stop.[186] The Court did not care. Because the defendants sped off at an "unreasonable" speed, the officer had probable cause to believe that a traffic violation had occurred and that was the end of the inquiry.[187] As long as the

---

[181] Rachel A. Harmon, *The Problem of Policing,* 110 MICH. L. REV. 761, 785 (2012) ("[A]ccording to legal scholars, the Constitution continues to be the primary means for regulating the police.").
[182] *See* Washington v. Davis, 426 U.S. 229, 238–42 (1976).
[183] 517 U.S. 806 (1996).
[184] *See id.* at 813.
[185] *Id.* at 808–09.
[186] *Id.* at 809.
[187] *Id.* at 810.

officers "could have" stopped the car for a traffic violation, it was irrelevant why they actually stopped the car.[188] No matter that because the police could always find probable cause for a traffic violation, it would be trivial for officers to stop someone on account of his race.[189] While the Court "of course agree[d] with petitioners that the Constitution prohibits selective enforcement of the law based on considerations such as race," it held that petitioners should make their claims under the Equal Protection Clause.[190]

In light of *Whren*, "scholars have written off the Fourth Amendment as a basis for challenging racially motivated searches and seizures."[191] But unintentional data-driven discrimination complicates the picture. In a sense, *Whren* was not actually a case about race. It held that probable cause was to be measured by an objective standard and that subjective motivations did not factor in.[192] When the *Whren* Court mentioned race, it held that "the constitutional basis for objecting to *intentionally* discriminatory application of laws is the Equal Protection Clause, not the Fourth Amendment."[193] The problems of discrimination in data mining, however, are not those of motive, conscious or unconscious. When police rely on a machine for their suspicion detection, the officers using the program are not even being subconsciously racist.[194] The people creating the model—as opposed to those using it—are more directly responsible for the discriminatory outcome, but neither are they likely to be relying, even unconsciously, on racial stereotypes.[195]

The conventional wisdom is actually more dismissive than the doctrine. As Devon Carbado summarized: "[F]or purposes of

---

[188] *Id.* at 809 (citing United States v. Whren, 53 F.3d 371, 374–75 (D.C. Cir. 1995)). This could be contrasted with the "would have" justification—"whether a police officer, acting reasonably, would have made the stop for the reason given." *Id.* at 810.

[189] *Id.* at 812.

[190] *Id.* at 813.

[191] Thompson, *supra* note 39, at 960–61.

[192] *Whren*, 517 U.S. at 813.

[193] *Id.* (emphasis added).

[194] *See* Barocas & Selbst, *supra* note 19, at 698–700.

[195] *Cf. id.* at 699 ("For example, the person who came up with the idea for Street Bump ultimately devised a system that suffers from reporting bias, but it was not because he or she was implicitly employing some racial stereotype. Rather, it was simply inattentiveness to problems with the sampling frame." (footnote omitted)).

Fourth Amendment law, race does not matter."[196] Kevin Johnson explained: "It may seem surprising to most readers, but the use of racial profiling by law enforcement authorities in the United States has long been permitted and encouraged, if not expressly authorized, by U.S. constitutional law."[197]

The commentary surrounding *Whren* has contributed to this perception. *Whren* was decided in the midst of then-emerging social science and cognitive psychology research into unconscious bias.[198] Humans categorize and stereotype in order to more quickly process information, and racial stereotypes are no different.[199] Anthony Thompson summarized this phenomena as follows: "As the human mind seeks to understand conduct, it looks to salient cues, such as race and ethnicity, and then draws on culturally embedded understandings to evaluate behavior."[200] The unconscious bias research demonstrates that no individual police officer could separate his thoughts about what looks like probable cause from his views about the correlations between racial, cultural, and gender identity and criminality.[201] Techniques like

---

[196] Devon W. Carbado, *(E)racing the Fourth Amendment*, 100 MICH. L. REV. 946, 1033 (2002). Paul Butler was even more explicit: The Fourth Amendment is "a project by the Burger, Rehnquist, and Roberts Courts to expand the power of the police against people of color." Paul Butler, *The White Fourth Amendment*, 43 TEX. TECH L. REV. 245, 246 (2010).

[197] Kevin R. Johnson, *How Racial Profiling in America Became the Law of the Land:* United States v. Brignoni-Ponce *and* Whren v. United States *and the Need for Truly Rebellious Lawyering*, 98 GEO. L.J. 1005, 1006 (2010).

[198] *See* Thompson, *supra* note 39, at 987–88 (describing research published almost concurrently with the *Whren* decision demonstrating that "negative attitudes towards African Americans create a perceptual norm of viewing African Americans as more prone to criminal conduct").

[199] *See id.* at 983.

[200] *Id.* at 983–84.

[201] *See id.* at 986–87 (noting that "police officers often proceed on the basis of 'traits' that, they assert, correlate with criminal behavior"). Discrimination scholarship around the same time was very focused on unconscious bias. Charles Lawrence, Linda Krieger, and many others were all examining the effects of unconscious bias and the Implicit Association Test (IAT) for anti-discrimination law. *See, e.g.*, Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161, 1177 (1995) (arguing that Title VII jurisprudence is inadequate to address unconscious bias); Charles R. Lawrence III, *The Id, the Ego, and Equal Protection: Reckoning with Unconscious Racism*, 39 STAN. L. REV. 317, 323 (1987) (explaining that an individual's unconscious biases "seem part of the individual's rational ordering of her perceptions of the world"). As Christine Jolls and Cass Sunstein observed, "[t]he scholarly literature critiquing existing antidiscrimination law, both constitutional and statutory, for its general failure to address the problem of implicit bias is voluminous." Christine Jolls & Cass R. Sunstein, *The Law of Implicit Bias*, 94 CALIF. L. REV. 969, 978

the Implicit Association Test and other reaction-time based instruments have demonstrated that even people who vehemently believe they are anti-racist exhibit unconscious biases,[202] including an association between race and criminality.[203] Because all humans exhibit unconscious bias, so too will all police officers.[204] The fact of unconscious bias is well enough understood that the use of seemingly neutral technology to take the decisions out of human hands is seen as a good thing.[205] This is why data mining is often sold as a way to reduce disparate outcomes.

After *Whren*, scholars such as Anthony Thompson argued that the case was the culmination of a broad, mistaken turn toward colorblindness in Fourth Amendment jurisprudence.[206] Thompson traced a history of Supreme Court decisions, beginning with *Terry v. Ohio*,[207] which constructed the narrative of a neutral, experienced police officer, "unaffected by considerations of race and who could be trusted even in a race-laden case like *Terry* to be acting on the basis of legitimate indicia of criminal activity."[208] In *Terry*, Detective McFadden's elephant-in-the-room testimony that

---

n.45 (2006) (collecting sources); *see also* Linda Hamilton Krieger & Susan T. Fiske, *Behavioral Realism in Employment Discrimination Law: Implicit Bias and Disparate Treatment*, 94 CALIF. L. REV. 997, 1003 n.21 (2006) (collecting sources).

[202] *See* Jerry Kang & Kristin Lane, *Seeing Through Colorblindness: Implicit Bias and the Law*, 58 UCLA L. REV. 465, 473 (2010).

[203] *See* Justin D. Levinson et al., *Guilty by Implicit Racial Bias: The Guilty/Not Guilty Implicit Association Test*, 8 OHIO ST. J. CRIM. L. 187, 190 (2010) (finding that subjects of a study held implicit associations between black people and the status of being guilty).

[204] *See* Tracey G. Gove, *Implicit Bias and Law Enforcement*, THE POLICE CHIEF, Oct. 2011, at 44, 50 ("Police officers are human and, as the theory contends, may be affected by implicit biases just as any other individual. In other words, well-intentioned officers who err may do so not as a result of intentional discrimination, but because they have what has been proffered as widespread human biases.").

[205] *See, e.g.*, Ellen Huet, *Rise of the Bias Busters: How Unconscious Bias Became Silicon Valley's Newest Target*, FORBES (Nov. 2, 2015), https://www.forbes.com/sites/ellenhuet/2015/11/02/rise-of-the-bias-busters-how-unconscious-bias-became-silicon-valleys-newest-target/.

[206] Thompson, *supra* note 39, at 981; *see also, e.g.*, DAVID COLE, NO EQUAL JUSTICE: RACE AND CLASS IN THE AMERICAN CRIMINAL JUSTICE SYSTEM 43 (1999); Lenese C. Herbert, *Bête Noire: How Race-Based Policing Threatens National Security*, 9 MICH. J. RACE & L. 149, 180 (2003) ("The Court has chosen . . . to view official action under the Fourth Amendment colorblind eye, side-stepping the pervasiveness of law enforcement that is race-based."). *See generally* Ian F. Haney López, *Post-Racial Racism: Racial Stratification and Mass Incarceration in the Age of Obama*, 98 CALIF. L. REV. 1023, 1064–66 (2010) (criticizing colorblindness in the criminal justice system more generally).

[207] 392 U.S. 1 (1968).

[208] Thompson, *supra* note 39, at 971.

he was "unable to say precisely"[209] what drew his eye to the black defendants became "a highly skilled officer's instinctive assessment that something in the situation seemed awry and worthy of investigation."[210] As Thompson explained, "[s]uch narratives permit the judges to clarify the events in their own minds and to present the facts and law in a manner that the legal community will generally accept."[211]

Before the rise of data-driven decision-making, conscious and unconscious human biases were the only possible sources of discrimination at the point of the decision. Thus, stating that the Fourth Amendment did not consider race was a clean and accurate shorthand, and it has become the dominant understanding.[212] This makes it highly unlikely, if theoretically possible, that a sympathetic judge will dig deep into the doctrine to find a Fourth Amendment violation due to disparate impact in suspect-based or person-based predictive policing. This low likelihood is compounded by the unavailability of disparate impact in equal protection doctrine, which suggests that the Constitution as a whole is not amenable to disparate impact as a theory.

2. Individualized Suspicion as a Disparate Impact Substitute. Arguably, the Fourth Amendment does not handle discrimination well because race was never meant to be its core concern. The Fourth Amendment does, however, address one form of racial bias—straightforward racial profiling. It can do so because of its individualized suspicion requirement.[213] Under the Fourth Amendment, police must have probable cause to effect a search or seizure, which includes a requirement that the cause be tied specifically to the person searched or things or people seized.[214] As

---

[209] *Terry*, 392 U.S. at 5 (quoting the detective's testimony).

[210] Thompson, *supra* note 39, at 969.

[211] *Id.* at 968–69 (footnotes omitted).

[212] *See id.* at 973 (arguing that "the Court's treatment of racial motivation" in *Terry* "established a pattern that would continue in the Court's subsequent Fourth Amendment cases").

[213] *See, e.g.*, Chandler v. Miller, 520 U.S. 305, 308 (1997) (stating that the Fourth Amendment "generally bars officials from undertaking a search or seizure absent individualized suspicion").

[214] Maryland v. Pringle, 540 U.S. 366, 371 (2003) (probable cause requires "a reasonable ground for belief of guilt . . . and that the belief of guilt must be particularized with respect to the person to be searched or seized" (internal quotation marks omitted) (citing Ybarra v. Illinois, 444 U.S. 85, 91 (1979))).

Andrew Taslitz put it, "individualized suspicion is the beating heart that gives probable cause its vitality."[215]

In the case of strict racial profiling, the individualized suspicion requirement serves as a proxy to prevent racially biased stops and searches. A person that is stopped on account of her race is not stopped for any reason demonstrating suspicion particular to her. But given the intentional nature of a racially profiled stop, it looks more like disparate treatment than disparate impact.[216] In racial profiling, race is the single—or dominant—factor in a stop whereas, if a suspect is stopped due to a data mining system with discriminatory outcomes, race will potentially be one of many factors (whether used implicitly or explicitly).[217]

Although case law hints that individualized suspicion could be repurposed to address the unique challenges posed by predictive policing, it will not work. But it is useful to understand why. Racial profiling is merely a special case of non-individualization that judges are attuned to. The requirements for a data mining system to become more individualized, and thus satisfy the Fourth Amendment, are orthogonal to discrimination; increases in the individualization of a model may unpredictably increase or decrease the disparate impact of the output.

a. Race and Individualization in the Case Law. The Supreme Court has drawn connections between individualization and discrimination, noting in the employment discrimination context "that Title VII requires employers to treat their employees as *individuals*, not 'as simply components of a racial, religious, sexual, or national class.' "[218] One major reason the Fourth Amendment requires individualized suspicion, and rejects broad surveillance, is to prevent the unbridled discretion that would allow for discrimination.[219]

---

[215] Andrew E. Taslitz, *What Is Probable Cause, and Why Should We Care?: The Costs, Benefits, and Meaning of Individualized Suspicion*, 73 LAW & CONTEMP. PROBS., Summer 2010, at 145, 145; *see also id.* (" 'Individualized suspicion,' the United States Supreme Court has suggested, is perhaps *the* most important of the four components of probable cause.").

[216] *See, e.g.*, Green v. Brennan, 136 S. Ct. 1769, 1784 (2016) (discriminatory intent is the "defining element" of disparate treatment).

[217] *See supra* Part II.C.

[218] Ariz. Governing Comm. v. Norris, 463 U.S. 1073, 1083 (1983) (quoting City of Los Angeles Dep't of Water & Power v. Manhart, 435 U.S. 702, 708 (1978)).

[219] *See* RANDALL KENNEDY, RACE, CRIME, AND THE LAW 5–6 (1997); Joh, *supra* note 159, at 28; *see also* United States v. Sokolow, 490 U.S. 1, 12 (1989) (Marshall, J., dissenting) ("By

In 2013, *Floyd v. City of New York*[220]—the famous stop-and-frisk case—demonstrated how a judge could treat a racial bias problem as a lack of individualized suspicion. *Floyd* involved a challenge to a New York Police Department program of routine stops and frisks, primarily in minority neighborhoods.[221] The case addressed a pattern of 4.4 million stops over an eight-year period, as well as nineteen stops of twelve individual plaintiffs.[222] Judge Scheindlin found that the program was racially biased and violated both the Fourth Amendment and the Equal Protection Clause of the Fourteenth Amendment.[223]

Importantly, Judge Scheindlin located the racial discrimination harms in the Fourteenth Amendment, but only found that the Fourth Amendment was violated due to the officers' repeated stops without individualized reasonable suspicion, even though the two violations stemmed from the same conduct.[224]

In the equal protection discussion, Judge Scheindlin found that the city intentionally discriminated, based on testimony that the policy required stopping "the right people," a term which was racially coded.[225] She also noted: "The NYPD's policy of targeting 'the right people' . . . is not directed toward the identification of a specific perpetrator. Rather, it is a policy of targeting expressly identified racial groups for stops *in general*."[226]

In the discussion of individual stops, Judge Scheindlin found several examples where reasonable suspicion was lacking because the only evidence was the race of the plaintiff and one or two other indicators that were not suggestive of criminality. Of one plaintiff,

---

requiring reasonable suspicion as a prerequisite to such seizures, the Fourth Amendment protects innocent persons from being subjected to 'overbearing or harassing police conduct carried out solely on the basis of imprecise stereotypes of what criminals look like, or on the basis of irrelevant personal characteristics such as race.'").

220 959 F. Supp. 2d 540 (S.D.N.Y. 2013).

221 *See id.* at 556.

222 *Id.* at 556, 561.

223 *Id.* at 562.

224 *Id.* at 562, 660–61 (noting that the "individual stop testimony corroborated much of the evidence about the NYPD's policies and practices"). Judge Scheindlin even began the discussion by quoting *Whren* for a statement that seems to mean the opposite of what people assume it stands for: "The Constitution prohibits selective enforcement of the law based on considerations such as race." *Id.* at 660 (quoting Whren v. United States, 517 U.S. 806, 813 (1996)).

225 *Id.* at 662–63.

226 *Id.* at 664 (footnote omitted).

she wrote: "Even if credited, Almonor's alleged furtive movements—looking over his shoulder and jaywalking—in combination with the generic description of young black male does not establish the requisite individualized suspicion that Almonor was engaged in criminal activity."[227] These findings about the individual plaintiffs demonstrate that, when police rely in large part on race to justify the suspicion necessary for a stop, the stop can violate the Fourth Amendment for being insufficiently individualized. Judge Scheindlin was able to find violations of the Equal Protection Clause in the individual cases as well, but consistent with the doctrine, they were an entirely separate discussion located outside of the Fourth Amendment.[228] In both holdings, the equal protection and individualization issues were clearly linked, but because of the structure of the Fourth Amendment, they had to be analyzed separately.

The Fourth Amendment's focus on individualization also explains treatment of race elsewhere in the doctrine. In two cases, *United States v. Brignoni-Ponce*[229] and *United States v. Martinez-Fuerte*,[230] the Supreme Court addressed "perceived Mexican ancestry" in the context of border searches.[231] The Court ruled that perceived Mexican ancestry could be a factor in the decision of whether to stop drivers "for brief inquiry into their residence status,"[232] but it could not be the only factor.[233] By stating that

---

[227] *Id.* at 630. Discussing another plaintiff, she noted that his

> "[P]resence in an area of expected criminal activity, standing alone, is not enough to support a reasonable, particularized suspicion that the person is committing a crime." . . . This, combined with the vague description of "black males" and the entirely unsuspicious act of putting one's hands in one's pockets in the wintertime, is a far cry from the *individualized* suspicion of wrongdoing that constitutes reasonable suspicion. Absent any other justification, there was no basis for a *Terry* stop, and there was certainly no basis to believe that McDonald was armed and dangerous.

Id. at 632–33 (quoting Illinois v. Wardlow, 528 U.S. 119, 124 (2000)).

[228] *See, e.g.*, *id.* at 633. Judge Scheindlin arguably could have gone further with data analysis than she did, but she did not have to because she was able to rely on damaging statements by the police department itself, a kind of evidence not available in most cases. Sharad Goel et al., *Combatting Police Discrimination in the Age of Big Data*, 20 NEW CRIM. L. REV. 181, 187 (2017).

[229] 422 U.S. 873 (1975).

[230] 428 U.S. 543 (1976).

[231] *See Brignoni-Ponce*, 422 U.S. 874–75; *Martinez-Fuerte*, 428 U.S. at 545.

[232] *Martinez-Fuerte*, 428 U.S. at 555.

race could not be the only factor, the Court removed the possibility of pure racial profiling and returned the question to a totality-of-the-circumstances determination of individualized suspicion.[234] This comports with the separate and widespread intuition that it would be absurd to exclude a suspect's race from visual descriptions.[235] Race as an identifier is acceptable if it sits alongside other factors. Accordingly, though "[c]ontesting the existence of reasonable articulable suspicion is a roundabout way of challenging police discrimination,"[236] it is worth asking if it is a viable one.[237]

But despite these links, individualization and discrimination are not tightly interconnected concepts. Disparate impact does not concern *individual* treatment, but rather it concerns *unfair* treatment on account of class membership.[238] If police cast more suspicion on someone because that person is of a certain race, that is unfair, but is it any less individualized than other observations? To answer that question it is important to have a fuller account of individualization in general, which is a surprisingly slippery concept.

---

[233] *Brignoni-Ponce*, 422 U.S. at 886–87. That these cases could have been considered instances of disparate treatment rather than disparate impact merely underscores the point that individualized suspicion has traction where race does not.

[234] *Cf.* Michael Tonry, *Legal and Ethical Issues in the Prediction of Recidivism*, 26 FED. SENT'G REP. 167, 170 (2014) ("Race, ethnicity, and religion are not to my knowledge anywhere used as an explicit factor in prediction instruments or in sentencing or parole policies. However, the use of any of them likely would be upheld, as it was in the profiling cases, so long as it was only one among several factors.").

[235] *See, e.g.*, David A. Harris, *Using Race or Ethnicity as a Factor in Assessing the Reasonableness of Fourth Amendment Activity: Description, Yes; Prediction, No*, 73 MISS. L.J. 423, 449 (2003) ("[R]ace is one of the most important physical characteristics of a criminal that one could include in this description . . . . Such an unchangeable, highly visible trait has real value in accurately describing the suspect.").

[236] Goel et al., *supra* note 228, at 196.

[237] A side benefit of the focus on individualization is that *Whren* is not implicated because probable cause itself is challenged. *Whren* only states that where there is an otherwise valid reason for the stop, an additional invalid reason cannot matter. Whren v. United States, 517 U.S. 806, 811 (1996). But under the individualization theory, the valid reason does not exist.

[238] *See, e.g.*, Samuel R. Bagenstos, *The Structural Turn and the Limits of Antidiscrimination Law*, 94 CALIF. L. REV. 1, 40–41 (2006) (discussing anti-discrimination law in terms of "unfairness"); George O. Luce, *Why Disparate Impact Claims Should Not Be Allowed Under the Federal Employer Provisions of the ADEA*, 99 NW. U. L. REV. 437, 440 (2004) ("[D]isparate impact theory was created as a form of strict liability that targets unfair *results*, without regard to the employer's motivation or intent.").

b. The Real Meaning of Individualization. Individualized suspicion is at the core of the Fourth Amendment's probable cause requirement. As the Supreme Court has described the constraint, "belief of guilt must be particularized with respect to the person to be searched or seized."[239] But that high-level statement hides a great deal of confusion and disagreement about both the practical meaning of individualization and the normative rationale for it.[240]

Intuitively, there seems to be a binary distinction between individualized and generalized observations. Generalized observations, including stereotypes, apply to a great number of people. For example: "People who wear red jackets are probably part of a gang." Conversely, individualized observations concern a single person: "He visited a known drug den." But when the police identify a person and stop him, the line between individual and general suddenly disappears.

Consider observations about Alice and Bob. Alice is seen wearing a red jacket. Bob is seen visiting a known drug den. The police stop both. Alice was stopped because the police know that

---

[239] Maryland v. Pringle, 540 U.S. 366, 371 (2003) (citing Ybarra v. Illinois, 444 U.S. 85, 91 (1979)).

[240] *See, e.g.*, Jane Bambauer, *Hassle*, 113 MICH. L. REV. 461, 464 (2015) ("The purpose of individualization is to minimize hassle. *Hassle* is the chance the police will stop or search an innocent person against his will."); *see also* Crespo, *supra* note 28, at 2102 (noting that "what *is* probable cause?" is one of the foundational constitutional criminal law questions the Court has left unanswered for decades). Many scholars have attempted to address the question, but few have done so satisfyingly. Several related questions have arisen around individualization. For example, can or should probable cause be quantified? *Compare* Orin Kerr, *Why Courts Should Not Quantify Probable Cause*, *in* THE POLITICAL HEART OF CRIMINAL PROCEDURE 112 (Michael Klarman et al. eds., 2012) (arguing against quantifying probable cause), *with* Erica Goldberg, *Getting Beyond Intuition in the Probable Cause Inquiry*, 17 LEWIS & CLARK L. REV. 789, 790–91 (2013) (arguing in favor of quantifying probable cause in light of police officers' increased reliance on technology that itself relies on probability and quantification). Can machines do what human police officers do? *See generally* Kerr, *supra*; *see also* Rich, *supra* note 157, at 897 (noting that Automated Suspicion Algorithms "are fundamentally incapable" of engaging in a totality-of-the-circumstances analysis); Taslitz, *supra* note 215, at 167 ("A computer-like set of 'if-then' rules for all police conduct is neither feasible nor wise."); Kiel Brennan-Marquez, *"Plausible Cause": Explanatory Standards in the Age of Powerful Machines*, 70 VAND. L. REV. 1249, 1298–1300 (2017) (arguing that machines cannot do the robust analytical work that humans do). *But see generally* Andrew D. Selbst, *A Mild Defense of Our New Machine Overlords*, 70 VAND. L. REV. EN BANC 87 (2017) (arguing that machines can provide some of the explanations that humans do, and should not be dismissed). Ultimately, if the answer to either of the last two questions is "no," then predictive policing based on data mining should simply be outlawed under the Fourth Amendment.

wearing a red jacket is a sign of being involved in gang activity. In order to connect the generalized statement about gangs to a particular individual, the officers made an observation about that individual—that Alice was wearing a red jacket. This piece of information is individualized in the exact same sense as Bob's observed choice to walk into a drug den—both relied on information about the individual. The complete syllogism is: (1) Alice was wearing a red jacket; (2) people who wear red jackets are probably part of a gang; and (3) therefore, Alice is probably part of gang. The police compared an observation about Alice to a known fact about the world, and then deduced that Alice was likely involved in criminal activity.

Bob's stop demonstrates the symmetry in reasoning. Simply noticing that Bob walked into a known drug den does not say anything about whether Bob was committing a crime. An extra step is needed and, in this case, it is just hidden: (1) Bob visited a known drug den; (2) *people who visit drug dens are probably involved in criminal activity*; and (3) therefore, Bob is probably involved in criminal activity.

In reality, neither of these factual scenarios is likely enough to stop Alice or Bob without more information.[241] There is nothing suggesting Alice is doing anything wrong at the moment—it is inconceivable that wearing a red jacket is, alone, enough to indicate criminal activity—and Bob could have been delivering a pizza. But what's important here is that the structure of the reasoning in both cases is identical. In both cases, police observed an innocent fact about a person, compared it to a general fact that connects the observed characteristic with crime, and used the comparison to add suspicion to that person. Both fact patterns used individualized and generalized information, and neither could have made any headway in assessing likelihood of criminal activity without reference to both. The basic structure of this reasoning applies in every case. Thus, there is no such thing as a truly individualized decision.[242] As Jane Bambauer puts it: "Cases

---

[241] *See* Ferguson, *supra* note 27, at 388 ("Knowing someone is a 'drug dealer' does not mean that the individual is actively dealing drugs at the moment of observation.").

[242] *See* CHRISTOPHER SLOBOGIN, PRIVACY AT RISK: THE NEW GOVERNMENT SURVEILLANCE AND THE FOURTH AMENDMENT 40 (2007) ("[T]he distinction between individualized and generalized suspicion is, in all relevant respects, meaningless."). The two scholars most

can be unique in the sense that they involve one-of-a-kind combinations of factors, but the *reasoning* of a case cannot be unique."[243]

This observation should not, however, be surprising. It is the very reason data mining works. Data mining operates on the understanding that a data subject's likely outcome for some query is similar to those people with whom the subject shares relevant attributes.[244] To discover something about a person, a data miner compares that person to everyone else that is like him in some specified way.[245] A data miner can add variables and make the model more accurate, but the fundamentals of the process remain unchanged.

Suppose that instead of the connection between red jackets and gangs, the police department had a model that predicted the likelihood of criminality based on jacket color, hairstyle, neighborhood, web surfing habits, and credit profile. Call it Model *M*. The reasoning surrounding Alice's stop would look the same. Now the syllogism is as follows:

> (1) Alice wore a red jacket, has short hair, lives in a certain neighborhood, primarily browses the internet

---

diametrically opposed on the importance of individual determinations, Andrew Taslitz and Frederick Schauer, believe, respectively, that humans are unique and should be identified as such, Taslitz, *supra* note 215, at 158–59, and that only aggregated generalizations matter, SCHAUER, *supra* note 142, at 68–69, 106 ("[I]ndividualized analysis is simply an aggregate of stereotypes . . . ."). But even they ultimately agree that "reasoning without some generalizations is impossible" and that "there is a *spectrum* of relative generality versus specificity." Taslitz, *supra* note 215, at 161.

After realizing that truly individualized suspicion is a myth, several scholars have concluded that individualized analysis should be wholly abandoned. Christopher Slobogin has proposed a proportionality principle for the probability side of the question, arguing that because we cannot individualize, we need to tighten our rules about probabilistic requirements into better-defined tiers than just probable cause and reasonable suspicion, so as to further constrain search. Christopher Slobogin, *The World Without a Fourth Amendment*, 39 UCLA L. REV. 1, 68–75 (1991). Similarly, Bernard Harcourt and Tracey Meares have pivoted from the lack of pure individualization to argue for randomized searches akin to checkpoints as the very basis for reasonableness in search. Bernard Harcourt & Tracey Meares, *Randomization and the Fourth Amendment*, 78 U. CHI. L. REV. 809, 850 (2011). Whether they are ultimately correct or not, individualization remains a central part of the case law, so it is worth exploring further.

[243] Bambauer, *supra* note 240, at 471.

[244] *See* Fayyad, *supra* note 84.

[245] *See id.*

> after midnight on her computer, and has a middling credit score; (2) people with red jackets and short hair that live in Alice's neighborhood, web surf after midnight, and have middling credit scores are often in gangs; and (3) therefore, Alice is probably in a gang.

This is much closer to what a predictive policing algorithm would look like because it uses multiple variables. Nevertheless, the addition of more variables did not fundamentally change the syllogism: (1) Person *X* is a member of Set *S*; (2) people in Set *S* probably have trait *T*; and (3) therefore, Person *X* probably has trait *T*.[246]

c. Individualization Will Not Prevent Disparate Impact. It is now possible to see the important difference between racial profiling and discriminatory predictive policing. The syllogism required for pure racial profiling takes the form: (1) Charlie is black; (2) black people are more likely to be criminals; (3) therefore, Charlie is more likely to be a criminal. No judge can permit a search based on this syllogism because of the second step—the racist generalization. But data mining adds more factors, potentially many more. Where the description of the subject includes more factors, the automatic rejection of the generalization step dissipates. Instead, the analysis looks like *Brignoni-Ponce* and *Martinez-Fuerte*, where Mexican heritage could permissibly be one factor, but not the only one.[247]

The conceptual clash goes even deeper. Just because nothing can be perfectly individualized does not mean that individualization is itself a meaningless concept. It is possible to think of individualization as a spectrum from "smaller generalizations" to "larger ones," where smaller generalizations are more individualized.[248] As Andrew Taslitz has observed, for the purposes

---

[246] *See* K.A. Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 COLUM. SCI. & TECH. L. REV. 2, 22 (2003).

[247] *See* United States v. Brignoni-Ponce, 422 U.S. 873, 886–87 (1975) ("The likelihood that any given person of Mexican ancestry is an alien is high enough to make Mexican appearance a relevant factor, but standing alone it does not justify stopping all Mexican-Americans to ask if they are aliens."); United States v. Martinez-Fuerte, 428 U.S. 543, 563 n.16 (1976) (citing a statistic suggesting only a small percentage of motorists of Mexican ancestry are stopped at the border to support the proposition that race is not the only factor used in making stops).

[248] *See* SCHAUER, *supra* note 142, at 68–69; Taslitz, *supra* note 215, at 157.

of the Fourth Amendment, the "practical question is where, as a matter of wise policy, to place ourselves on [the] continuum" from generalized to individualized decisionmaking.[249] In the context of data mining, while it may not be possible to define "individualized" in an absolute way, it is still coherent to speak of a model that is *more or less* individualized than another. In those terms, if a model is individualized enough, it satisfies the Fourth Amendment. For the purposes of discrimination, the question is whether increasing individualization in order to satisfy the Fourth Amendment will ameliorate or exacerbate discriminatory outcomes.

In a data mining model, a more individualized system will be able to make determinations—whether positive or negative—about fewer people overall. Thus, there are two aspects of a trained data mining model that can individualize it. More features could be observed, or the data collected could be more granular.[250] Adding features is the difference between "red jacket" Alice and "Model *M*" Alice. The set of people who look like Alice in Model *M*—that is, wear red jackets, have short hair, live in Alice's neighborhood, web surf after midnight, and have middling credit scores—is a subset of the people who wear red jackets. By including more features, a model can return a prediction that is keyed more specifically to an individual. It is a "smaller generalization."[251] The same is true of more granular data. For example, if the location data exists at the

[249] Taslitz, *supra* note 215, at 161.

[250] While additional or more granular features in training data individualize a model, these narrowing criteria are only useful if the information about a suspect includes enough data to take advantage of the additional comparison points. If "black male" is all that is entered into the algorithm, the algorithm will output an average result for all black males, even if the model has been trained on more specific information. Thus, a third practical requirement for more individualized decisions is to use more available data about the subject.

[251] *See* SCHAUER, *supra* note 142, at 68. There is one important caveat. Each of the additional features attached to the suspect must actually add to the suspicion. If red jackets indicate gang membership at the same rate as someone that matches Alice's profile in Model *M*, then none of the other factors in Model *M* are doing any work. And if red jacket-based suspicion is only a New York phenomenon, then the location is no longer doing any work. If the entire determination can be explained by a subset of the total variables, then the determination is only as individualized as the greatest number of people that the subset of variables applies to.

neighborhood level, rather than the city level, fewer people are implicated in any result, whether positive or negative.[252]

This formulation explains the acceptance of using race as an identifying factor, as long as it is only one of several factors.[253] In the cases where race adds suspicion, it is in a sense no different than identifying someone by their red jacket from the perspective of the Fourth Amendment's individualized suspicion standard. If a person is identified by race in a physical description, it becomes relevant. If a person is of "apparent Mexican ancestry" near the border, as in *Martinez-Fuerte*,[254] it makes him more likely to be present illegally, as a matter of sheer statistics. But race is a factor that applies to all people, most of them innocent. Accordingly, race alone is insufficient to individualize the inquiry, and more features are necessary.

Thus, adding features or making them more granular makes the model more individualized. Those new features, however, will suffer from the same biases of the features that make up the model in the first place.[255] The data corresponding to those features could be of lower quality, unrepresentative, or mislabeled, and thus skew the output to be more or less discriminatory. Similar effects can be observed with respect to data granularity. Going from city, to neighborhood, to hotspot or "high crime area" could either increase or decrease disparate impact.[256] Thus, adding features and increasing granularity is just as likely to exacerbate

---

[252] This mathematical perspective on individualization also explains much of the case law, which requires adding variables in the human systems of observation. As Jane Bambauer notes: "When assessing an officer's decision to stop or search somebody, courts prefer to receive a long list of reasons justifying the decision. The more reasons the agent can recount, the better." Bambauer*, supra* note 240, at 496. This is because "[a]dding factors to the Venn diagram has an exclusionary effect" and "[c]ourts are reassured by longer lists of justifications because these lists roughly signal that the agent's model cannot scale to a large number of people, many of whom may be innocent." *Id.* at 497.

[253] *See* United States v. Brignoni-Ponce, 422 U.S. 873, 886–87 (1975).

[254] 428 U.S. 543, 563 (1976).

[255] *See supra* Part II.B.

[256] *See* Adam Benforado, *The Geography of Criminal Law*, 31 CARDOZO L. REV. 823, 846–48 (2010) ("Grow up on the wrong side of town, and your chances of ending up in the back of a squad car increase dramatically."); Ferguson, *supra* note 76, at 217 (describing the correlation between "high-crime areas," "low income communities," and "communities of color"); David A. Harris, *Factors for Reasonable Suspicion: When Black and Poor Means Stopped and Frisked*, 69 IND. L.J. 659, 660–61, 676–78 (1994) (explaining that *Terry* stops are applied "disproportionately to poor, to African Americans, and to Hispanic Americans" because these groups are the most likely to live in "high crime areas").

disparate impact as it is to reduce it, and there is no way to tell beforehand whether increased individualization will lead to better or worse outcomes for members of various protected classes.[257]

In sum, the Fourth Amendment will not provide a regulatory solution. According to many scholars, the Fourth Amendment as a regulatory tool is mistakenly colorblind at best, and outright racist at worst.[258] Claims of disparate impact will likely fall on the deaf ears of judges who intrinsically understand the Fourth Amendment this way. Individualization will not serve as a substitute for a disparate impact claim because increasing individualization in data mining models is at cross-purposes with the reduction of disparate impact, and racial profiling is truly a special case.

B. THE FAILURE OF THE ANTI-DISCRIMINATION LIABILITY MODEL

Lacking a constitutional response, an answer must come in some form of legislation. Because the problem is disparate impact, the obvious approach would be a law modeled on existing anti-discrimination laws, such as Title VII. But the discrimination in data mining systems confounds the typical model of anti-discrimination regulation. Barocas and I explained why this is the case for Title VII in a prior work,[259] but here I expand the argument to laws modeled on an anti-discrimination framework

---

[257] Reintroducing humans to the process does not solve the problem either. Ric Simmons and Michael Rich have both suggested that, while data mining should be part of the overall suspicion-generation process, it should only be used in conjunction with human policing. *See* Rich, *supra* note 157, at 983 ("[A] person trained in making individualized suspicion determinations must be the final assessor of the totality-of-the-circumstances, including both the ASA's prediction and any other relevant available data . . . ."); Simmons, *supra* note 52, at 991 ("A human being at least has the potential to incorporate new observations, but a predictive algorithm is limited by its previous programming."). In these proposals, predictive algorithms become similar to any other direct observations of suspicious behavior, after which the police use their discretion to decide whether to act. But such a process only serves to double the sources of disparate impact. Not only will the effects of the algorithms' disparate impact go unrecognized by police and be treated as a neutral fact, but the discretion that the "neutral algorithm" is supposed to solve again becomes a part of the overall decision.

[258] *See, e.g.*, Butler, *supra* note 196, at 246; Carbado, *supra* note 196, at 967–68; David A. Harris, *"Driving While Black" and All Other Traffic Offenses: The Supreme Court and Pretextual Traffic Stops*, 87 J. CRIM. L. & CRIMINOLOGY 544, 550–53 (1997); Maclin, *supra* note 38, at 338.

[259] Barocas & Selbst, *supra* note 19, at 715.

more generally. The lessons from this regulatory strategy's failure will, in turn, inform the features a new regulatory scheme should possess.

1. Beyond Title VII. There are three parts to a disparate impact case. First, the plaintiff must make a prima facie showing of disproportionate impact from a defendant's decision on a protected class.[260] Second, if the plaintiff succeeds, then the burden shifts to the defendant to make a showing of "business necessity."[261] In Title VII, this essentially amounts to a showing that the decision is "job related,"[262] but it can be generalized to a requirement of fitness for the purpose—i.e., whether the decision is truly related to the legitimate outcome sought. Third, if the defense succeeds, the plaintiff may then demonstrate that there was an equally effective, but less discriminatory, tool available that the decision-maker refused to use.[263] In Title VII, this is referred to as the "alternative employment practice" test.[264] Thus, the disparate impact injury can be described as the use of a decision mechanism with a disproportionate impact on a protected class when it is not a good predictor of future outcomes or there is an alternative, equally effective, and less discriminatory decision mechanism available.

In analyzing a disparate impact case based on predictive policing, the first part of the test—that there is an observable disparate impact—should be assumed to be true, or there is no reason for the discussion. And in our prior work, Barocas and I demonstrated why the business necessity defense will usually be satisfied by data mining.[265] Essentially, data mining is a powerful statistical prediction method for the legitimate outcomes sought by the decision-maker, so as long as it is done well enough, it will satisfy the test.[266] Thus, the only remaining question is whether

---

[260] *See* 42 U.S.C. § 2000e-2(k) (2012).

[261] *See id.*

[262] *See* Barocas & Selbst, *supra* note 19, at 705 (noting that "all circuits seem to accept varying levels of job-relatedness").

[263] *Cf.* 42 U.S.C. § 2000e-2(k)(1)(A)(ii).

[264] *Id.*

[265] Barocas & Selbst, *supra* note 19, at 706–09.

[266] *See id.* at 708–09 ("[T]here is good reason to believe that any or all of the data mining models predicated on legitimately job-related traits pass muster under the business necessity defense.").

there is an equally effective, but less discriminatory, alternative. The obvious answer is to fix the model. But that is deceptively difficult.[267]

First, as discussed in Part II, data miners must make necessarily subjective judgments that are unavoidable aspects of system design. How the target variable is defined may lead to more or less discriminatory results, yet it will not necessarily be obvious which choice will lead to worse results on protected classes.[268] As Oscar Gandy has noted, "certain kind[s] of biases are inherent in the selection of the goals or objective functions that automated systems will [be] designed to support."[269] Should predictions be broken down by type of crime, or some other value? Are three different crime categories better than seven? It is impossible to say in the abstract. The police officers will choose based on their views of how to most faithfully and usefully predict and prevent crime within their jurisdictions.[270] That is their primary—and arguably only—goal. But the decisions will necessarily have different effects on different racial, ethnic, and socioeconomic groups.

Similarly, for the number and granularity of features chosen, one might think that the optimal goal is greater precision. But police resources for data collection are not unlimited, and some data that would be theoretically ideal may not be available due to either cost or inaccessibility. The choice of how much or which data to collect will often be made on the basis of cost or convenience, and thoughts about what is even possible will depend on the available data, rather than on imagining what is possible if someone goes out and collects more.[271] This risk is magnified by the availability of third-party data collected for the commercial market,[272] where the features chosen come as a package deal, and the overall package is determined not by what is best for the particular police jurisdiction, but by which data broker has the best price. Like target variable definition, the decisions about data

---

[267] *See id.* at 709–11.
[268] *See id.* at 715 (discussing the difficulties of defining the target variable).
[269] Gandy, *supra* note 108, at 39.
[270] *See* Bachner, *supra* note 8, at 20.
[271] *See* Calders & Žliobaitė, *supra* note 155, at 52–53.
[272] *See* PERRY ET AL., *supra* note 9, at 77 (explaining that many police departments "subscribe to commercial services" for crime data).

collection and costs of more comprehensive data collection are subjective, and the "correct" ones are impossible to identify in the abstract. In both cases, however, decisions made on the basis of "better" policing will invariably have different effects on vulnerable groups. There is no system of disparate impact liability that will hold decision-makers liable for decisions in service of the proper goals where there is no clear less discriminatory alternative.[273]

Once a discriminatory result is discovered, the reasons for it are likely to be unclear. Even if the result can be traced to a data quality problem, those problems are often quite complicated to rectify. It might be easy to determine that something is off about the data, but it can be more difficult to figure out what that something is. While the source of some biases might be clear on the face of the analysis, most others are not. For example, the potential for skewed data because of biases in distribution of prior police resources is the most well-known data quality issue with predictive policing.[274] Thus, it makes sense to check if that is what might be causing the skew. When past crime data is involved in the calculus, then the source of bias is clear.[275] But commercial or social media data is likely to have biases that are not as apparent and could be skewing the program's outputs in unknown ways, with no way to investigate.

Even if all the sources of bias are identified, the magnitude of each source's effect is still likely unknown. Suppose a department implements Chicago's "heat list," with the intention of naming the top 400 most likely people to be involved in a shooting.[276] Then suppose that 95% of the names on the list are those of black men. By almost any measure, this hypothetical list appears to disproportionately target black men. But there is no baseline by which to measure *how much* the bias is contributing. It is not obvious how to subtract out what portion of the 95% prediction is due to discriminatory results, and what is attributable to black men in the city actually being more likely to be involved in violent

---

[273] *Cf.* 42 U.S.C. § 2000e-2(k)(1)(A) (2012).
[274] *See supra* notes 127–31 and accompanying text.
[275] *See supra* notes 127–31 and accompanying text.
[276] *See supra* note 17 and accompanying text.

crime than their population numbers might suggest.[277] But without an omniscient outside source of data, it is unclear where such baseline data should come from.

The problem will not always be data quality, either. Returning to the hypothetical "heat list" that is 95% populated by black men, suppose—ignoring all the problems just discussed—it were possible to determine that the prediction should have been 45% black men, with the other 50% resulting from a discriminatory algorithm. If someone wanted to correct the disparate impact, how would he go about doing so? The first step would be to identify the source of disparate impact. Bad data is only one of the possible ways the model could have gone wrong. It could also be the problem definition, feature choice, or feature granularity. But just like determining the source of bad data, figuring out whether the real culprit is the data or something else may not even be possible. Worse, if there is both bad data and other sources of bias, the *degree* to which each is at fault is uncertain. If bad data is always an option, it will not be clear on the face of the problem if the data miner has chosen a target variable or particular features that lead to the disparate impact, and, therefore, it will not be clear what should be fixed.

In other words, trying to determine the single cause of disparate impact that results from these machines is a flawed exercise, and often practically impossible. With data mining, *some* disparate impact will occur as a result of the use of the system in the first place, unless it is specifically mitigated.[278] The only way to ensure that a data mining system will not discriminate is to not use it.[279] When there are background differences in the arrest rates of protected and unprotected classes, it will be impossible to have a system that can remove disparate impact by all available measures.[280]

---

[277] *See* Barocas & Selbst, *supra* note 19, at 718.

[278] Michael Feldman et al., *Certifying and Removing Disparate Impact*, PROC. 21ST ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 259 (2015).

[279] *See* Barocas & Selbst, *supra* note 19, at 728. This might seem like an argument to entirely block their use, but other systems reliably result in discriminatory outcomes as well. The ones resulting from data mining are just more easily quantified.

[280] *See* Kleinberg et al., *supra* note 75 (concluding that, if base rates of a predicted characteristic differ, it is impossible to satisfy all of the identified "fairness constraints" simultaneously in predicting that trait).

Disparate impact doctrine considers the presence of discrimination a binary question and either finds someone at fault for causing it or finds no discrimination at all. Perhaps counterintuitively, this is true of both disparate treatment and disparate impact regimes. In a disparate treatment regime, the human decision is obvious—it is the intentional decision to discriminate.[281] In a disparate impact regime, however, the law still attempts to trace a human decision to the injury—specifically, the choice of the screening mechanism that resulted in the disparate outcome. A finding of disparate impact liability is still a finding of blameworthiness, and ultimately, once we accept that data mining is permissible, there is nobody to blame for the *presence* of disparate results. Therefore, disparate impact doctrine will blame no one.

2. Disagreements About the Meaning of Discrimination. The failures of the liability model described above presume that its normative goals are well understood and agreed upon. In most contexts where discrimination is an issue, however, normative agreement does not exist, and predictive policing adds an additional wrinkle. If one could imagine a data mining scenario in which all the variables mentioned so far are correct—ignoring that "correct" is itself subjective—the disparate impact might still reflect the reality according to the thing one is trying to predict.[282] To the extent that, because of background inequalities, criminal behavior is truly distributed unequally between racial groups,[283] the proper response is up for debate.

Consider again the "heat list" hypothetical. If the "accurate" percentage of the list that should be black is 45%, and if the actual population is 33% black (Chicago, according to the 2010 census),[284] then the discrepancy reflects criminal disparities between races.

---

[281] *See, e.g.*, St. Mary's Honor Ctr. v. Hicks, 509 U.S. 502, 507 (1993).

[282] *See* Barocas & Selbst, *supra* note 19, at 691–92.

[283] *See* Benforado, *supra* note 256, at 847–48 (" '[N]eighbourhood effects on offending can and do sometimes exist,' and in what amounts to a vicious circle, these effects may be perpetuated by the offending that they encourage." (footnote omitted) (quoting Anthony E. Bottoms, *Place, Space, Crime, and Disorder*, *in* THE OXFORD HANDBOOK OF CRIMINOLOGY 528, 559 (Mike Maguire et al. eds., 4th ed. 2007))).

[284] U.S. CENSUS BUREAU, *Profile of General Population and Housing Characteristics: 2010*, https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk (last visited Sept. 21, 2017).

But people will disagree about whether that 12% difference is a discriminatory result as the law understands discrimination. On the one hand, the model is saying to focus on a list that is disproportionately made up of black men. On the other, if black men are disproportionately committing the crimes, then the machine cannot be blamed for picking up on that.

This debate reflects a longstanding tension between anti-discrimination law's two competing normative models: anticlassification and antisubordination.[285] Under the anticlassification theory, the responsibility of the law is simply to prevent harm to protected classes due to the choices of decision-makers.[286] The anticlassification model would suggest that, if it is not the predictive policing tool itself that is causing the discrimination, then the users of the tool cannot be held responsible for it. By contrast, antisubordination sees the project of antidiscrimination as one designed to eliminate substantive inequality as a result of membership in the protected class, no matter the cause.[287] Under the antisubordination theory, racially disparate results are simply intolerable and should be rectified.

There is a valid argument to be made for each perspective. A focus on antisubordination suggests that, to the extent that data mining reflects the status quo, addressing disparate impact amounts to asking police to catch fewer criminals and to predict fewer crimes. There is also an inherent unfairness in making a single decision-maker responsible for all of the structural discrimination in society, as discussed above. At the same time, no matter the cause, the criminal justice system's focus on communities of color "lock[s] people of color into a permanent second-class citizenship."[288] Even if the algorithm is accurate, merely accepting that result seems intolerable from the perspective of working to fix the inequities reflected in the disparate—yet accurate—result.[289]

---

[285] *See* Helen Norton, *The Supreme Court's Post-Racial Turn Towards a Zero-Sum Understanding of Equality*, 52 WM. & MARY L. REV. 197, 206–15 (2010); *see also* Bagenstos, *supra* note 238, at 40–42 & nn.214–15 (2006) (collecting sources).

[286] *See* Norton, *supra* note 285, at 208–09.

[287] *See id.* at 206.

[288] ALEXANDER, *supra* note 99, at 13.

[289] Further complicating data-driven discrimination in the policing context, even before a sector in which an anti-discrimination law is in place, is that the "discrimination" is not

Importantly, though, this difficulty might be more theoretical than real. As mentioned in Part II, it may be impossible to tell *when* the disparate impact truly reflects reality.[290] The "accurate" 45% number is made up for this hypothetical, but it is often not detectable in reality for the reasons laid out above. In most cases, "reducing the disparate impact will necessitate open-ended exploration without any way of knowing when analysts have exhausted the possibility for improvement."[291] Bad crime data makes this a fairly likely scenario, and it will often be impossible to fully disentangle disparate impact due to algorithm design. If there is uncertainty whether the result reflects reality, then the arguments for removing disparate impact and the difficulties of doing so revert to the prior discussion.

In a wrinkle specific to predictive policing, what counts as adverse may not be as clear as a context such as employment. If a model recommends that the police focus resources disproportionately on communities of color, then, as this Article argues, the additional focus is likely to lead to a discriminatory harm by bringing more people of color into contact with the police, which will lead to disproportionate arrest rates in those communities. But what about the other residents of the community? If there is more crime in a neighborhood (either perceived or real), the residents there may actually want a greater police presence. Deciding to forego the added police presence could even be a disparate treatment violation in itself.

This normative impasse may be unresolvable. Instead, it suggests that democratic input could be valuable, especially

---

necessarily tethered to protected class, because there is no statute that says it must be. In one sense of the word, data mining discriminates by its very nature. It sorts and selects between otherwise similarly situated people to find the major points along which predictions can be made. *See* Barocas & Selbst, *supra* note 19, at 677. In this sense, there is nothing inherently special about legally defined protected classes. Data mining will indiscriminately disadvantage the disadvantaged along whatever axis is relevant. *See generally* Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014). As a society, we have decided that such discrimination is more concerning when that axis is a protected class. But perhaps in the policing context, it is as damaging to overpolice poor neighborhoods as it is communities of color. For the purposes of this proposal, discrimination is treated in the traditional sense, but in the face of data mining-focused discrimination, that decision need not be fixed.

290 *See supra* Part II.C.

291 *See supra* Part II.C.

engagement with smaller constituencies in towns and neighborhoods.[292] In the absence of a more definitive answer, localized political groups are more likely to agree outright or compromise about the right approach to issues that affect them locally. It will not always work, but resolution is more likely to come at the local level than from a top-down approach, and then knowledge can be shared between the smaller subgroups.[293] Thus, the approach proposed in Part IV attempts to take this into account.

IV. TOWARD A REGULATORY SOLUTION

A new legislative approach is necessary. The last Part offered some hints as to what a new approach should look like, but a legislative approach must address a few more considerations. At a high level, there are actually two separate problems to address: (1) the substantive challenge—predictive policing systems have the potential for discriminatory results; and (2) the precursor problem—a lack of knowledge about the effectiveness and discriminatory impact of predictive policing systems. Currently, police do not understand the technology they adopt. They do not understand or investigate either the technology's effectiveness or its side effects.[294] It is hard to say in the abstract what stronger regulatory solutions may be required, or how big a problem the technology poses in reality, until more information about the technology's implementation is created. And right now, there is little incentive for the various actors to understand it at all.

Therefore, this Part argues that before adopting predictive policing technology, police should be required to create "algorithmic impact statements" (AISs), modeled on the environmental impact statements (EISs) of the National Environmental Policy Act (NEPA).[295] Impact statements have

---

[292] *See* Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 317–18 (1998) (arguing that local civil engagement is effective in improving police services).

[293] *See id.* at 326 (describing the benefits of "reciprocating consultation between central and local levels").

[294] Ferguson, *supra* note 12, at 1117 ("Whether good, bad, ineffective, or distracting, the long-term trend has been to adopt predictive technologies regardless of effectiveness.").

[295] *See* 42 U.S.C. § 4332(C) (2012).

become a much-emulated regulatory tool where the problem at hand is a lack of knowledge about the effects of a particular type of decision.[296] While AISs will not necessarily achieve the full measure of accountability that will eventually be required, they will be useful—and perhaps necessary—to determine what, if anything, society will need to do next.

A. ALGORITHMIC IMPACT STATEMENTS

The goal of this proposed legislation is not necessarily to curtail the use of new predictive policing technologies. The AISs would have two purposes. First, they would ensure "that the agency, in reaching its decision, will have available, and will carefully consider, detailed information concerning significant [discriminatory] impacts."[297] Second, they would "guarantee that the relevant information will be made available to the larger audience that may also play a role in both the decisionmaking process and the implementation of that decision."[298] The EIS model is important precisely because, without the requirement, "information about prospective . . . harms and potential mitigating measures" simply would not exist.[299] Presently, there is too little public engagement regarding predictive policing, and other users of algorithmic tools do not currently know what the ultimate social effects of these tools will be. Without some intervention, critical knowledge about these systems might never be created.

Since the passage of NEPA in 1969, impact assessments have been a model for "action-forcing" regulation designed to push decision-makers to do their homework and engage with the public.[300] As Bradley Karkkainen has observed: "NEPA is without question the most widely emulated of the major U.S. environmental laws. It has inspired dozens of 'little NEPAs' at the

---

[296] *See* Bradley C. Karkkainen, *Toward a Smarter NEPA: Monitoring and Managing Government's Environmental Performance*, 102 COLUM. L. REV. 903, 905 (2002).

[297] *See* Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 349 (1989).

[298] *Id.*

[299] Winter v. Nat. Res. Def. Council, Inc., 555 U.S. 7, 23 (2008).

[300] *See e.g.*, Kleppe v. Sierra Club, 427 U.S. 390, 409 (1976) ("Section 102(2)(C) is one of the 'action-forcing' provisions intended as a directive to 'all agencies to assure consideration of the environmental impact of their actions in decisionmaking.' " (quoting Conference Report on NEPA, 115 CONG. REC. 40416 (1969))).

state and local levels . . . and countless imitators in other fields."[301] And the relevance of environmental law to this problem may be more than just coincidence. Cathy O'Neil has argued that "[i]f you think of [data mining] as a factory, unfairness is the black stuff belching out of the smoke stacks. It's an emission, a toxic one."[302] Thinking of discrimination as an inevitable byproduct of data mining's machinery can be a useful construct. Others have also drawn analogies between data and environmental concerns. Michael Froomkin has written that lessons from NEPA can be applied to the regulation of mass surveillance.[303] Similarly, Dennis Hirsch has likened data breaches to oil spills in service of an argument that other parts of environmental law can be a model for greater data privacy protection as well.[304]

Some of the existing NEPA-inspired legislation bears directly on issues of data use, discrimination, or both. For example, administrative agencies are required to conduct privacy impact assessments (PIAs) "when developing or procuring information technology systems that include personally identifiable information."[305] Other countries use them too. The United Kingdom recommends "equality impact assessments" as part of the "public sector Equality Duty," which requires attention to discrimination concerns in all activities by public bodies.[306] And the European Union's General Data Protection Regulation (GDPR) and accompanying Police and Criminal Justice Authorities Directive both require data protection impact assessments (DPIAs) whenever data processing "is likely to result in a high risk to the

---

[301] Karkkainen, *supra* note 296, at 905–06.

[302] O'NEIL, *supra* note 111, at 95.

[303] *See generally* A. Michael Froomkin, *Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements*, 2015 U. ILL. L. REV. 1713, 1755. In a similar vein, David Wright and Charles Raab have argued for "surveillance impact assessments" (SIAs) to address the effects of surveillance systems. The SIA process would be "similar to that of a privacy impact assessment (PIA), but [ ] an SIA must take account of a wider range of issues, impacts and stakeholders." David Wright & Charles D. Raab, *Constructing a Surveillance Impact Assessment*, 28 COMPUTER L. & SECURITY REV. 613, 613 (2012).

[304] *See* Dennis D. Hirsch, *The Glass House Effect: Big Data, the New Oil, and the Power of Analogy*, 66 ME. L. REV. 373, 375 (2014).

[305] Kenneth A. Bamberger & Deirdre K. Mulligan, *Privacy Decisionmaking in Administrative Agencies*, 75 U. CHI. L. REV. 75, 76 (2008).

[306] *Review of Public Sector Equality Duty*, GOV.UK, https://www.gov.uk/government/groups/review-of-public-sector-equality-duty-steering-group (last visited Sept. 22, 2017).

rights and freedoms of natural persons."[307] The recitals in both laws suggest that discrimination is one such high-risk concern.[308] Outside of the data context, several states have adopted racial impact assessments as part of sentencing policy to learn more about how proposed changes are likely to affect people of color before implementation.[309]

The NEPA process and larger impact statement model have drawn strong support and sharp criticism. In the benefits column, they can force government agencies to both think hard about the collateral effects of the proposed policy and justify the policy to the public.[310] Scholars looking back have argued that the process has been ingrained into the core of how administrative agencies now do their work,[311] and it has "create[d] powerful pressures on agency decisionmakers to avoid the most environmentally damaging courses of action."[312] As for the negatives, the NEPA process has been criticized as overly long, too costly, and ultimately toothless.[313] Much of the criticism, though, is about the particulars of NEPA, rather than the core principles—that it is too

---

[307] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, art. 35, 2016 O.J. (L 119 1) 1, 59 [hereinafter General Data Protection Regulation or GDPR]; Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of the Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and on the Free Movement of Such Data, and Repealing Council Framework Decision 2008/977/JHA, art. 27 O.J. (L 119) 89, 116 [hereinafter Police and Criminal Justice Authorities Directive].

[308] GDPR, *supra* note 307, Recital 75, at 15; Police and Criminal Justice Authorities Directive, *supra* note 307 Recital 51, at 96–97.

[309] *See* Catherine London, *Racial Impact Statements: A Proactive Approach to Addressing Racial Disparities in Prison Populations*, 29 L. & INEQ. 211, 226–31 (2011) (noting that "[r]acial impact statements are modeled on fiscal and environmental impact statements").

[310] Paul J. Culhane, *NEPA's Impacts on Federal Agencies, Anticipated and Unanticipated*, 20 ENVTL. L. 681, 690 (1990) ("The resources policy community should not overlook the importance of a simple, but definitely non-trivial outcome of NEPA. Environmental impacts *are* now considered in making natural resources decisions."); SERGE TAYLOR, MAKING BUREAUCRACIES THINK: THE ENVIRONMENTAL IMPACT STATEMENT STRATEGY OF ADMINISTRATIVE REFORM 251 (1984).

[311] *See* TAYLOR, *supra* note 310, at 262 ("Since the advent of NEPA, environmental concerns have been officially incorporated into every agency's charter.").

[312] Karkkainen, *supra* note 296, at 905.

[313] *See id.* (summarizing the criticisms of the EIS process).

easy for an agency to avoid doing an EIS,[314] that the explanation requirements are not backed up by accountability for failure to act on them,[315] or that judges have been too deferential to agency factfinding.[316] The fact that an EIS is prescriptive and frozen in time is another commonly-cited weakness of NEPA.[317] Thus, the criticisms of NEPA can be lessons for future implementations, while the core rationales for the EIS model fit exceptionally well with the considerations laid out in the prior section.

1. AIS Requirements. Before discussing costs and benefits of the model, it is important to understand what such legislation would actually entail, beginning with the requirements of the AIS itself. The regulation at the "heart of the environmental impact statement"[318] lists its six requirements:

> (a) Rigorously explore and objectively evaluate all reasonable alternatives, and for alternatives which were eliminated from detailed study, briefly discuss the reasons for their having been eliminated.
> (b) Devote substantial treatment to each alternative considered in detail including the proposed action so that reviewers may evaluate their comparative merits.

---

[314] *See id.* at 908 (explaining how, through a "mitigated Finding of No Significant Impact . . . an agency might avoid the cost and administrative burden associated with a full [EIS]").

[315] Matthew J. Lindstrom, *Procedures Without Purpose: The Withering Away of the National Environmental Policy Act's Substantive Law*, 20 J. LAND RESOURCES & ENVTL. L. 245, 246 (2000) ("Now in its thirtieth year, the Supreme Court and disinterested executive leadership have essentially rendered NEPA's substantive objectives and declarations for environmental quality legally impotent."); *see also* Ray Clark, *NEPA: The Rational Approach to Change*, *in* ENVIRONMENTAL POLICY AND NEPA: PAST, PRESENT AND FUTURE 15, 23 (Ray Clark & Larry Canter eds., 1997) (observing that many public officials regard NEPA as merely "a rigid paperwork exercise" rather than "a way to maintain or achieve environmental objectives").

[316] *See* Aliza M. Cohen, Note, *NEPA in the Hot Seat: A Proposal for an Office of Environmental Analysis*, 44 U. MICH. J.L. REFORM 169, 198 (2010).

[317] *See* COUNCIL ON ENVTL. QUALITY, THE NATIONAL ENVIRONMENTAL POLICY ACT: A STUDY OF ITS EFFECTIVENESS AFTER TWENTY-FIVE YEARS 31–33 (1997) (recommending adaptive management in the face of high ecosystem uncertainty); J.B. Ruhl & Robert L. Fischman, *Adaptive Management in the Courts*, 95 MINN. L. REV. 424, 429–30 (2010) ("[F]rom the earliest emergence of ecosystem management policy, there has been broad consensus among resource managers and academics that adaptive management is the only practical way to implement ecosystem management.").

[318] 40 C.F.R. § 1502.14 (2017).

> (c) Include reasonable alternatives not within the jurisdiction of the lead agency.
> (d) Include the alternative of no action.
> (e) Identify the agency's preferred alternative or alternatives, if one or more exists, in the draft statement and identify such alternative in the final statement unless another law prohibits the expression of such a preference.
> (f) Include appropriate mitigation measures not already included in the proposed action or alternatives.[319]

Five of these requirements are directly portable to AISs, and the other suggests a similar rule that AISs can adopt. Each are considered here in turn.

a. Rigorously Explore and Objectively Evaluate All Reasonable Alternatives. The twin primary purposes of an AIS are (1) that police departments (and potentially other agencies) think hard about and investigate the particular choices they make rather than blindly using the first algorithm they think of or encounter, and (2) that they create the knowledge regarding the ultimate effects of their choices. Thus, the requirement to rigorously explore and objectively evaluate the algorithms is central. Exploration and evaluation does not necessarily mean that the algorithms themselves need be human-readable or interpretable[320]—such a requirement could potentially hinder the adoption of a wide swath of possible software.[321] Rather, the data miners must (1) explain the various design choices, (2) measure the resulting efficacy using the best available audit methods, and (3) evaluate the resulting disparate impact for the various systems and configurations. In this regime, it is not necessary to know exactly *why* those differences occur—the why might not even be

---

[319] *Id.*

[320] *See generally* Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 662–72 (2017) (arguing that technical tools can be designed for automatically ensuring regularity).

[321] *See* Zachary C. Lipton, *The Mythos of Model Interpretability*, PROC. 2016 ICML WORKSHOP ON HUMAN INTERPRETABILITY IN MACHINE LEARNING 96 ("[A]rguments against *black-box* algorithms appear to preclude any model that could match or surpass our abilities on complex tasks.").

answerable—but in order to evaluate competing alternatives, the efficacy and side effects must be measured.

The word "reasonable" does a lot of work here and will require interpretation.[322] There are essentially an infinite number of possible alternative algorithms: new target variables can be chosen, new or better data collected or purchased, and new features examined. The number of alternatives considered will depend on the specifics of the technologies at issue, and will mostly be left to the considered judgment of the professionals. Designers should also consider state-of-the-art "brute force" disparate impact removal. In other words, they should consider pre- or post-processing techniques to remove disparate impact and conduct an analysis of whether, and to what degree, it can be done without affecting the overall quality of the algorithm. As with any reliance on reasonableness, this will spur litigation.[323] Ultimately, there is no way around this. Regulating police in this way is a somewhat radical change that some departments will resist. As such, the threat of litigation for unreasonably skimpy AISs is a necessary element.

b. Devote Substantial Treatment to Each Alternative. An AIS should "[d]evote substantial treatment to each alternative considered in detail . . . so that reviewers may evaluate their comparative merits."[324] This requirement is self-explanatory. The information in an AIS "must be of high quality."[325] Reviewers need enough detail to know whether the department actually considered the various alternatives, or whether they should challenge the decision procedurally.

---

[322] *Cf.* Nat. Res. Def. Council, Inc. v. Morton, 458 F.2d 827, 837 (D.C. Cir. 1972) ("[R]easonable alternatives does not require 'crystal ball' inquiry. Mere administrative difficulty does not interpose such flexibility . . . as to undercut the duty of compliance 'to the fullest extent possible.' But . . . [t]he statute must be construed . . . not to demand what is, fairly speaking, not meaningfully possible . . . ." (footnote omitted)). When the Supreme Court narrowed the scope of alternatives needing consideration in *Vermont Yankee Nuclear Power Corp. v. Natural Resources Defense Council, Inc.*, 435 U.S. 519 (1978), it held that not every "conceivable" alternative need be considered, but failed to attach the word "reasonable" to alternatives anywhere in the opinion. *Id.* at 551.

[323] *See, e.g.*, *Morton*, 458 F.2d at 829 ("This appeal raises a question as to the scope of the requirement of the National Environmental Policy Act (NEPA) that environmental impact statements contain a discussion of alternatives." (footnote omitted)).

[324] 40 C.F.R. § 1502.14(b) (2017).

[325] 40 C.F.R. § 1500.1(b) (2017).

c. Include Reasonable Alternatives Not Within the Jurisdiction. NEPA requires that an EIS "[i]nclude reasonable alternatives not within the jurisdiction of the lead agency."[326] This is the requirement that does not directly apply to an AIS, at least as envisioned here.[327] The NEPA process applies to any agency that might take an action that affects the environment. Multiple agencies, such as Fish and Wildlife Service, National Forest Service, Bureau of Land Management, etc. could end up coordinating, and a modification that might mitigate the proposed impact might be outside their reach—either in the jurisdiction of another agency, or outside what congressional authorization permits them to do.

The jurisdictional limits on police are few. Usually they are charged broadly with enforcing the criminal law.[328] The decision to use predictive policing tools is not one that will be considered by more than one type of agency, so the requirement is mostly inapplicable here. But if considered for applications outside policing, this requirement may once again apply. For example, if a particular data set would help alleviate some discrimination that would otherwise result, but that data set is owned and used by another agency, then the AIS should consider it. This may be true even where law prevents inter-agency access. As the Council on Environmental Quality has explained, an "EIS may serve as the basis for modifying the Congressional approval or funding in light of NEPA's goals and policies."[329] In a similar vein, the AIS might spur political action locally, either to grant or restrict funding, or to bar certain choices by the police department. Because the AIS

---

[326] 40 C.F.R. § 1502.14(c) (2017).

[327] It would apply directly if the AIS requirement were expanded to any public sector use of data mining, which, while not necessarily a large logical leap, is outside the scope of this Article.

[328] *See* Friedman & Ponomarenko, *supra* note 28, at 1831, 1844.

[329] COUNCIL ON ENVTL. QUALITY, EXEC. OFFICE OF THE PRESIDENT, FORTY MOST ASKED QUESTIONS CONCERNING CEQ'S NATIONAL ENVIRONMENT POLICY ACT REGULATIONS § 2(b) (1981), https://energy.gov/sites/prod/files/G-CEQ-40Questions.pdf; *see also* Nat. Res. Def. Council, Inc. v. Morton, 458 F.2d 827, 837 (D.C. Cir. 1972) ("The mere fact that an alternative requires legislative implementation does not automatically establish it as beyond the domain of what is required for discussion . . . .").

may generate the only source of information about these algorithms,[330] this part of the AIS is also potentially useful.

d. Include the Alternative of No Action. The NEPA regulations specifically require that the "alternative of no action" be included.[331] This is crucial here as well. If the disparate impact is unavoidable and of an unacceptable degree, then the police must actively consider not going forward with their proposal and declining to adopt the technology. Predictable consequences of the "no action" alternative, such as potential limitations on criminal enforcement, should also be included in the analysis.[332]

The "no action" alternative is particularly important as the primary mechanism by which police departments will ask whether adoption of the predictive policing system will be better or worse than the discriminatory status quo. There will be a great temptation to give short shrift to this section. Police (as with many other institutional actors) will often seek reasons to adopt new technology, rather than reasons not to adopt technology that may have harmful effects.[333] And often these systems are sold as cost-cutting measures for cash-strapped departments for whom the budget is much more tangible than potential externalities that affect the community.[334] But if there is a risk that these systems will result in net harm, then the "no action" alternative must be preserved.

This is another reason that the AIS must be performed before a project proceeds, rather than as it is underway. Once money and other resources are committed, the question of whether the project goes forward at all instead becomes a question of how it will proceed. After all, if the AIS says that it should not go forward, then the person that already committed the resources to it—who is

---

[330] *Cf.* Winter v. Nat. Res. Def. Council, Inc., 555 U.S. 7, 23 (2008) ("[W]ithout [an EIS], there may be little if any information about prospective environmental harms and potential mitigating measures.").

[331] 40 C.F.R. § 1502.14(d) (2017).

[332] *Cf.* COUNCIL ON ENVTL. QUALITY, *supra* note 329, § 3.

[333] *See generally* Barry Friedman, *We Spend $100 Billion On Policing. We Have No Idea What Works*, WASH. POST (Mar. 10, 2017), https://www.washingtonpost.com/posteverything/wp/2017/03/10/we-spend-100-billion-on-policing-we-have-no-idea-what-works/?utm_term=.0ec3fde479cb.

[334] *See, e.g.*, *The ROI of Predictive Policing*, PREDPOL, http://info.predpol.com/roi-of-predictive-policing (last visited Sept. 23, 2017) (advertising the purported efficiency-maximizing benefits of the PredPol system).

also likely involved in the AIS process—will have wasted a whole lot of money. The path dependence and political pull of demonstrable progress all but guarantee that if an AIS is not performed prior to the beginning of the project, the "no action" alternative will not be considered as fully as required.[335]

e. Identify the Agency's Preferred Alternative in Draft Assessment. In the EIS model, public comments are solicited between the production of a draft and the final EIS.[336] The same would be true of AISs. Additionally, a police department must indicate its preferred alternative among the algorithm design choices within the draft AIS.[337] This is important so that reviewers and commenters know both what to focus on and what the likeliest course of action will be if nothing changes.

f. Include Appropriate Mitigation Measures. Lastly, the AIS should include mitigation measures.[338] That is, if the police determine that they need to use these models and recognize that some disparate impact will result, they must explore ways of mitigating the impact within the harmed communities. Broadly, these could be interventions other than arrests for identified people.[339] One possibility has already been discussed. Chicago originally planned to use the Strategic Subjects List to call in social services to aid those that were likely to be involved in crime. Instead, Chicago used the list to identify people for greater surveillance.[340] The plan, in theory, would have been a good example of mitigation. Another mitigation strategy could be combining location-based predictive policing with community policing strategies that are proven to reduce fear and improve community relations with police.[341] Mitigation procedures include

---

[335] *See* W. Brian Arthur, *Competing Technologies, Increasing Returns, and Lock-In by Historical Events*, 99 ECON. J. 116, 116 (1989) (explaining the "lock-in" effect of technology adoption).

[336] *See* 40 C.F.R. § 1503.1(a) (2017).

[337] *Cf.* 40 C.F.R. § 1502.14(e) (2017).

[338] *Cf.* 40 C.F.R. § 1502.14(f) (2017).

[339] *See* Harmon, *supra* note 177, at 359.

[340] *See supra* notes 171–72 and accompanying text.

[341] *See* Gary Cordner, *Community Policing*, *in* THE OXFORD HANDBOOK OF POLICE AND POLICING 148, 164 (Michael D. Reisig & Robert J. Kane eds., 2014) ("[C]ommunity policing seems to have clear-cut advantages over competing police strategies when it comes to making the public feel safer and enhancing the public's satisfaction with the police.").

anything that could lessen the burden that the use of predictive policing technology may create or exacerbate.

2. The Remaining AIS Framework. Though AISs are the heart of the proposed regulation, there are important details that must attend them to make the proposal effective. Particularly, there must be opportunities for public comment and strong judicial oversight, both procedural and substantive.

Public comment is an incredibly important part of the AIS process.[342] Under NEPA, regulations require two notice-and-comment periods during the EIS process: one to define the scope of the proposed draft EIS,[343] and one after the draft EIS to allow public comment on the information generated before the final EIS.[344] In the final version, agencies must respond to any opposing views that were not adequately addressed in the draft EIS.[345] These two comment periods can be ported directly into the AIS framework. The first would allow civil society groups and members of the public to register concerns with predictive policing that pertain to the specific jurisdiction as a general matter, and the second would allow public response to the direction that the department is heading and the concerns the department has already considered.

Public comment would not be a panacea. For example, where there are conflicting definitions of fairness,[346] the police department or another agency will have the initial power to determine which definition of fairness they think is the most proper. One thing public comment would accomplish, however, is to inform the agency that there *are* several standard definitions of

---

[342] *See* Jonathan Poisner, *A Civic Republican Perspective on the National Environmental Policy Act's Process for Citizen Participation*, 26 ENVTL. L. 53, 55 (1996) ("[C]itizen participation in the creation of NEPA-mandated [EISs] has, in all likelihood, spawned the largest amount of citizen participation in environmental decision making over the last two decades.").

[343] *See* 40 C.F.R. § 1501.7 (2017) ("As soon as practicable after its decision to prepare an [EIS] and before the scoping process the lead agency shall publish a notice of intent (§ 1508.22) . . . . (a) As part of the scoping process the lead agency shall: (1) Invite the participation of . . . other interested persons . . . .").

[344] *See* 40 C.F.R. § 1503.1 (2017).

[345] *See* 40 C.F.R. § 1502.9(b) (2017).

[346] *See* Julia Angwin & Jeff Larson, *Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say*, PROPUBLICA (Dec. 30, 2016), https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say.

fairness with different implications within that jurisdiction. Before the ProPublica article on Northpointe sparked a very public hearing on definitions of fairness,[347] it was not clear that the engineers designing the algorithm would have understood that they were encoding normative judgments into their system. Public comment, especially at the scoping stage, can bring issues like that to the fore early on. Once raised, at either the scoping or draft stage, the department would be required to discuss why it focused on a particular choice, and to include this concern as part of its publicly-reasoned decision.

Judicial review is also key to ensuring that the AIS process has some teeth.[348] The procedural requirements can be ported directly from the Administrative Procedure Act (or state equivalents), which requires courts to set aside arbitrary and capricious actions by agencies as well as actions that fail to comport with procedural requirements.[349]

A weakness of NEPA is that it is wholly procedural and lacks any substantive force. Under NEPA, once an environmental impact is identified, the agency is free to simply ignore the problem and forge ahead.[350] That is, an agency must strictly

---

[347] *See* Sam Corbett-Davies et al., *A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not That Clear*, WASH. POST (Oct. 17, 2016), https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-than-propublicas/ ("[A]t the heart of their disagreement is a subtle ethical question: What does it mean for an algorithm to be fair?"); *see also* Angwin et al., *supra* note 44 (investigating the racial bias of Northpointe's COMPAS program used to set bail and sentence duration in courts across the country).

[348] Judicial review has been quite important to NEPA. *See generally* Nicolas C. Yost & James W. Rubin, *Administrative Implementation of and Judicial Review Under the National Environmental Policy Act*, *in* 1 THE LAW OF ENVIRONMENTAL PROTECTION § 10:1, Westlaw (database updated Apr. 2017). *See also* Kleppe v. Sierra Club, 427 U.S. 390, 421 (1976) (Marshall, J., concurring in part and dissenting in part) ("[T]his vaguely worded statute seems designed to serve as no more than a catalyst for development of a 'common law' of NEPA. To date, the courts have responded in just that manner and have created such a 'common law.' Indeed, that development is the source of NEPA's success.").

[349] *See* 5 U.S.C. § 706 (2012).

[350] *See* Karkkainen, *supra* note 296, at 910–11 (noting that "NEPA itself leaves decisionmakers discretion to ignore" information provided in an EIS); *see also, e.g.*, Robertson v. Methow Valley Citizens Council, 490 U.S. 332, 350–52 (1989) (noting that because NEPA is procedural, the agency is not required to act on the EIS); *id.* at 351 n.16 ("NEPA merely prohibits uninformed—rather than unwise—agency action."); Strycker's Bay Neighborhood Council, Inc. v. Karlen, 444 U.S. 223, 227 (1980) (holding that NEPA is designed to generate informed decisionmaking but does not require the agency to elevate environmental concerns over other considerations).

adhere to the procedural requirements of NEPA, but a reviewing court cannot second-guess the choices the agency makes about how to value competing alternatives. Many commentators believe NEPA could have had substantive force, had it not been eviscerated by unfriendly courts.[351] In *Strycker's Bay Neighborhood Council, Inc. v. Karlen*, for example, the Supreme Court overturned the Second Circuit's holding that an agency's environmental determinations "should be given determinative weight."[352] The Court wrote that NEPA is "essentially procedural" and not subject to even arbitrary and capricious substantive review.[353]

An AIS-implementing statute need not be restricted by such a limitation. The AIS statute could expressly call for substantive review of the choices made in the AIS. NEPA arguably did so. As Nicholas Yost has argued, the legislators passing NEPA thought they were embedding substantive policy goals in the statute, and early courts recognized that purpose.[354] The AIS review could include something more searching than arbitrary and capricious review, such as "hard look" review,[355] which requires some

---

351 *See, e.g.*, Philip Michael Ferester, *Revitalizing the National Environmental Policy Act: Substantive Law Adaptations from NEPA's Progeny*, 16 HARV. ENVTL. L. REV. 207, 217 (1992) (Section heading: "NEPA Before the Supreme Court: Extinguishing Substantive Review."); David R. Hodas, *NEPA, Ecosystem Management and Environmental Accounting*, 14 NAT. RESOURCES & ENV'T, Winter 2000, at 185, 186–87 (arguing that the Supreme Court has been too deferential to agencies, thus undermining the substantive goals of NEPA); Lindstrom, *supra* note 315, at 249 ("NEPA's directives to federal agencies . . . [a]re not flowery sentiments . . . they are positive law binding on all parts of the federal government."); Nicholas C. Yost, *NEPA's Promise—Partially Fulfilled*, 20 ENVTL. L. 533, 534 (1990) (arguing that "[s]ubstantive review under NEPA" is "essentially unfulfilled"); *The National Environmental Policy Act: An Interview with William Hedeman, Jr.*, EPA J., Nov.–Dec. 1980, at 29, 30, https://archive.epa.gov/epa/aboutepa/national-environmental-policy-act-interview-william-hedeman-jr.html ("I feel that much of NEPA's problem in the past has been the manner in which it has been interpreted by the courts. . . . Unfortunately, most of this litigation has focused on procedural compliance with the requirements of NEPA rather than getting to the basic substantive mandates of the Congress as reflected in NEPA's goals and policies.").

352 444 U.S. at 227.

353 *Id.*

354 *See* Yost, *supra* note 351, at 535–39.

355 Kleppe v. Sierra Club, 427 U.S. 390, 410 n.21 (1976) (citing Nat. Res. Def. Council v. Morton, 458 F.2d 827, 838 (D.C. Cir. 1972)). Some state NEPA equivalents, like California's, have more searching review. *See* CAL. CODE REGS. tit. 14, § 15151 (2017) ("An EIR should be prepared with a sufficient degree of analysis to provide decisionmakers with

balancing of substantive costs and benefits.[356] If the AIS framework's purpose is clearly stated to be prevention of discrimination, and there is an equally clear statement that discrimination considerations deserve extra weight, judges could ensure that the procedural components were guided mostly by these concerns.[357]

There is a limit, though. The level of substantive review cannot be so high that a court is simply substituting its judgment for that of the agency.[358] This is true for a few reasons. The AIS relies on the expertise of the police in understanding the crime control needs of their jurisdictions, and that expertise must be accorded *some* deference.[359] But given the current state of extreme deference to police,[360] it is hard to imagine courts being willing to inject themselves much in police affairs, so this concern may not manifest without broader cultural and political changes. Second, there is still normative disagreement about what constitutes discrimination, when it occurs, and what definitions of fairness are appropriate. While the resulting disparate impact of different algorithm designs can be compared against each other in the AIS, in the end, substantive review of the AIS must compare the chosen version of the model against some standard. And there is no global normative standard by which to judge the adequacy of these decisions in an absolute sense.[361] Giving courts strong substantive review over implementation of the recommendations of AISs shifts

information which enables them to make a decision which intelligently takes account of environmental consequences.").

[356] Cass R. Sunstein, *Constitutionalism After the New Deal*, 101 HARV. L. REV. 421, 463–74 (1987) (discussing costs and benefits of the "hard-look doctrine").

[357] Philip Ferester has argued that courts ignored the substantive provisions of NEPA because they are "vaguely drafted aspirational commands and are far more difficult to apply" than the procedural provisions. Ferester, *supra* note 351, at 208. But he did not see this as an inevitable consequence of the model; rather, he expressly argued that NEPA should be amended based on state versions of NEPA with much stronger substantive provisions. *Id.* at 230.

[358] *See* Friedman & Ponomarenko, *supra* note 28, at 1876 ("Policing agencies today possess unfathomable discretion, the appropriate control for which is democratically sanctioned rules—not judicial judgments or (worse yet) naked deference.").

[359] *See id.* at 1840 ("[C]ourts hold out the promise of greater deference to agency interpretations of vague statutory terms when these interpretations are arrived at through more deliberative processes.").

[360] *Id.* at 1890, 1892.

[361] *See* Bagenstos, *supra* note 238, at 34–40 (discussing the "deeply controversial" nature of substantive discrimination standards).

the responsibility for that determination from a combination of policing agencies and public comment to the courts. On the whole, this may be desirable if police prove unwilling to respond to public concerns, but it must be understood that substantive review has this shifting effect. It is a careful balance, but one that, at present, suggests that substantive judicial review will help more than hurt.

B. ANTICIPATING OBJECTIONS

The AIS is not a perfect or complete solution. On its own terms, it is more concerned with information creation than a determination of when predictive policing becomes too discriminatory, and how to prevent that from occurring. This is intentional. It is difficult to make policy recommendations when the public knows so little about the technology and its implementation. Even acknowledging that, however, the proposed AIS regulation will encounter several objections, which this section attempts to anticipate and address.

1. Costs. A major criticism of the EIS model is its length and cost.[362] In the AIS case, the average police department is small and may not be able to absorb these costs.[363] There are at least three reasons, however, to think this might not be as big a barrier as it first appears.

First, to a degree, the cost is a feature rather than a bug. Police departments should not be adopting technology without considering the ramifications, and the most likely reason that they would do such a thing is the falling cost of technology. Society has seen what comes from the availability of free technology to police departments in the context of leftover or obsolete military equipment.[364] Police have been given billions of dollars in military equipment,[365] which in turn has led to increased violent confrontations and default deployment of SWAT teams for minor

---

[362] *See* Karkkainen, *supra* note 296, at 905; Sarah E. Light, *NEPA's Footprint: Information Disclosure as a Quasi-Carbon Tax on Agencies*, 87 TUL. L. REV. 511, 536 (2013).

[363] *See* DUREN BANKS ET AL., U.S. DEP'T OF JUSTICE, NCJ 249681, NATIONAL SOURCES OF LAW ENFORCEMENT EMPLOYMENT DATA 1 (2016) ("The most common type of agency is the small town police department that employs 10 or fewer officers.").

[364] *See* Alicia Parlapiano, *The Flow of Money and Equipment to Local Police*, N.Y. TIMES (Dec. 1, 2014), http://www.nytimes.com/interactive/2014/08/23/us/flow-of-money-and-equipment-to-local-police.html.

[365] *See id.*

drug raids and to break up protests.[366] Police surveillance has also exploded largely because of the low cost of technology.[367] The unthinking adoption of technology due to its low cost is precisely what this proposal aims to prevent, and aside from the substantive requirements of AISs, if the cost alone causes police departments to slow down and ask if they really need this new equipment, that will aid in this proposal's goals. Intentionally adding friction to decision-making is sometimes the correct policy choice.[368]

Second, small departments being unable to meet the cost of producing AISs is not as overwhelming a problem as it might appear. Police departments with ten or fewer officers are likely located in small towns without much need of predictive policing technology.[369] Arguably, the only reason they would consider such a tool is the low cost. Additionally, there would be nothing preventing this proposal from being adopted at the state level rather than the local level—in fact, that seems more likely. Once adopted, if state legislatures believe their towns should have predictive policing equipment, they can work in partnership with municipal police departments to produce AISs and defray the costs, rather than leave the AISs solely to the localities. The added benefit of such an arrangement is that the states themselves would provide an extra layer of consideration before the public has to even weigh in, allowing an additional opportunity to weigh the costs and benefits.

Third, the AIS can import some of the cost-saving features from NEPA. Because of the cost, agencies will go far out of their way to

---

[366] *See generally* RADLEY BALKO, RISE OF THE WARRIOR COP: THE MILITARIZATION OF AMERICA'S POLICE FORCES (2013) (noting the increasingly militarized nature of American police forces); Molly Redden, *The Pentagon Just Realized It Gave Too Much Military Equipment to the Ferguson Police*, MOTHER JONES (Aug. 12, 2015), http://www.motherjones.com/politics/2015/08/pentagon-forces-ferguson-return-two-humvees-police-militarization-program/.

[367] *See* Kevin S. Bankston & Ashkan Soltani, *Tiny Constables and the Cost of Surveillance: Making Cents Out of* United States v. Jones, 123 YALE L.J.F. 335, 335 (2014), http://www.yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones.

[368] *See* Paul Ohm & Jonathan Frankle, *Proof of Work: Learning from Computer Scientific Approaches to Desirable Inefficiency* (draft on file with author) (manuscript at 37-41) (discussing the role of friction in various contexts).

[369] *See* BANKS ET AL., *supra* note 363, at 1; Slobogin, *supra* note 25, at 135.

avoid having to create an EIS.[370] This can be done because the NEPA framework provides an agency with three options before it seeks to implement a proposed course of action, of which the EIS is the most demanding. The agency must either (1) apply a categorical exclusion,[371] (2) perform a smaller "environmental assessment"[372] (EA), which, if it results in a "finding of no significant impact," ends the process,[373] or (3) conduct an EIS. Categorical exclusions are activities pre-determined to have no significant environmental impact.[374] EAs allow federal agencies to avoid conducting a full-blown EIS. Therefore, the cost criticism is directly tied to separate criticism that agencies can too easily avoid conducting an EIS.[375]

But what does it mean to "avoid" an AIS in this way? Some argue that these avoidance mechanisms are the reason NEPA has been a shadow success.[376] If the reason that agencies can avoid EISs is because they can create situations that already have been approved, or take environmental concerns in early, then that is a benefit of the scheme, whether or not an EIS is the end result.

Allowing departments to avoid AISs could have the desirable effect of encouraging them to think about disparate impact early in the process, so as to avoid the extra cost later.[377] At the moment, police have no incentive to evaluate whether their systems are in fact discriminatory. At a high level, the entire point of this proposal is to get police to figure out how to do better on that score. If there were a standardized set of test suites that could demonstrate that a police department's chosen technology does not have a significant disparate impact, then all the better. A market

---

[370] *See* Light, *supra* note 362, at 535 ("The onerous EIS requirement creates incentives for agencies to employ mitigation measures to bring the impact of an action on the environment below the 'significance' threshold (that is, the point at which an EIS must be prepared).").

[371] 40 C.F.R. § 1508.4 (2017).

[372] 40 C.F.R. § 1508.9 (2017).

[373] 40 C.F.R. § 1508.13 (2017).

[374] *Id.*

[375] *See* Karkkainen, *supra* note 296, at 905 (explaining that because of "the high triggering threshold before an EIS is required," there are few actions that "receive full NEPA scrutiny").

[376] *See, e.g.*, TAYLOR, *supra* note 310, at 251.

[377] *Cf.* Karkkainen, *supra* note 296, at 913 ("[A]gencies now operate under strong internal and external pressures to select and design projects from the start with an eye toward reducing their adverse environmental consequences, precisely because they wish to avoid embarrassing NEPA disclosures.").

would undoubtedly develop in order to sell certified testing suites that police departments could use to test their technology. If avoiding an AIS is significantly less expensive, then police departments will be encouraged to design to a standard that can be easily and cheaply tested, while still flexible enough to be properly applied to their jurisdictions.

Additionally, if police departments can use discrimination-aware data mining[378] at the design stage to reduce the ultimate disparate impact, then that too could provide another way to avoid a full AIS. This is comparable to the "mitigated finding of no significant impact" (mitigated FONSI), a practice that federal agencies use to avoid EISs.[379] Although such mitigation is often criticized as undercutting the goals of the statute, it may actually accomplish them indirectly by forcing agencies to begin with a better design to avoid the costs associated with a full EIS.[380] If the way to avoid an AIS is to instead do better at avoiding disparate impact early in the process, then that is a perfectly acceptable victory.

Therefore, costs should not be as big a problem as people fear, and cost-cutting measures could be included in the model.

2. Ineffectiveness of the EIS Model. Another criticism of the impact statement model is its supposed ineffectiveness. This critique has a couple flavors. One is that impact statements are simply weak and do not accomplish anything.[381] This concern can be driven by fears of poor design or regulatory capture, among other things. It is also a reflection of differing ideas about what constitutes an impact statement. Another flavor, addressed above, is that agencies can too easily avoid doing a full AIS, and will do so if given the opportunity.

The first concern is borne of observed practice. In some contexts, impact statements are fundamentally different documents than those advocated here, and much less robust. Fiscal impact statements attach a dollar figure to legislative

---

[378] *See, e.g.*, Feldman et al., *supra* note 278, at 259.

[379] *See* Karkkainen, *supra* note 296, at 933–37 (detailing the use of mitigated FONSIs to avoid EISs).

[380] *See id.* at 935 ("Indeed, mitigated FONSIs might be considered evidence the NEPA is having a tangible, proactive, environmentally beneficial effect on agency decisionmaking, not dissimilar to the one NEPA's authors intended, albeit through an unexpected route.").

[381] *See supra* notes 315–17 and accompanying text.

proposals and perhaps provide some context about a state's budget.[382] The "Anticipated Surveillance Impact Report," now required in Santa Clara before police purchase surveillance equipment, is a short explanation of anticipated impact.[383] Several states have begun to require racial impact statements (RISs) in the sentencing context to collect data and determine disparate impact resulting from new sentencing policy.[384] While there is value in the transparency that these documents provide, these versions of impact statements are not as robust as the proposed AIS. They can run a mere handful of pages, whereas regulations were required to limit EISs to 150 pages, except in cases of "unusual scope or complexity," which are afforded 300.[385] They do not have the specific requirements of an EIS, "such as requiring community outreach through the use of comment periods or obligating decision-makers to seriously consider alternatives."[386] Instead, they often involve reporting on the impact of an already determined course of action. As a result, the reports are less informative. And because the only question posed is an up or down vote on a proposal,[387] which already has momentum behind it, these impact statements are less likely to lead to better alternatives. As proposed here, the AIS law would require something much more robust and consequential. If an AIS as written looked like these other impact statements, it would be a violation of the procedural requirements, which could then be enforced in court.

Another reason for the apparent ineffectiveness of impact statements may result from a mismatch between the goals of an AIS and what advocates truly want. As discussed above, AISs have two goals: forcing early consideration of different options and the resulting externalities, and knowledge creation. If the

---

[382] *See generally Preparation of Fiscal Analysis*, NAT'L COMM. STATE LEGISLATURES, Apr. 2002, http://www.ncsl.org/research/elections-and-campaigns/fiscal-impact-statements.aspx.

[383] *See* SANTA CLARA COUNTY, CAL. CODE div. A40, § 7(D) (2016).

[384] *See* Jessica Erickson, Comment, *Racial Impact Statements: Considering the Consequences of Racial Disproportionalities in the Criminal Justice System*, 89 WASH. L. REV. 1425, 1444–45 (2014); London, *supra* note 309, at 226–31.

[385] 40 C.F.R. § 1502.7 (2017).

[386] *See* Erickson, *supra* note 384, at 1428 (arguing that RIS legislation could be more effective if it resembled the EIS framework).

[387] *See, e.g.*, NAT'L COMM. STATE LEGISLATURES, *supra* note 382 (describing how fiscal impact statements are created only to aid voters in deciding how to vote on a given ballot measure).

predictive policing system at issue turns out to be highly discriminatory, AISs will not have the firepower to rectify the results. The AIS is designed to change the decision process, and failing that, it is fundamentally a transparency measure. Transparency, however, does not automatically lead to accountability,[388] so faulting a transparency statute for failing to provide full reform is a case of mismatched expectations.

The sentencing context once again provides such an example. Marc Mauer, Executive Director of the Sentencing Project and an early proponent of the RIS strategy in sentencing, has argued that, in addition to providing information, these statements could lead legislatures to actually reduce the disparities.[389] Acknowledging that such a desire was perhaps "wishful thinking,"[390] it was nevertheless clear that such reform was the intended goal of the proposal. Therefore, if disparities are not reduced, RISs could be seen as a failure.

By that metric, AISs may fail as well. But that is not the best metric by which to measure success. If the AISs demonstrate internally that predictive policing is flawed, then that will be an invisible victory even if the AIS never appears in that form. If, however, the AIS comes out and demonstrates that predictive policing is highly discriminatory, while there is a chance that this will cause such embarrassment that reform will result, there are also good reasons to think further political intervention will be required. This is especially true in light of the extreme emphasis on procedural goals in the existing interpretations of NEPA. If AISs are interpreted as purely procedural, they will lose even their low level of substantive impact. While this may be an inherent limitation of the impact statement model, it only suggests that the

---

[388] *See generally* Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, NEW MEDIA & SOC'Y 1 (2016). *See also* W.C. Bunting, *The Regulation of Sentencing Decisions: Why Information Disclosure Is Not Sufficient, and What to Do About It*, 70 N.Y.U. ANN. SURV. AM. L. 41, 60 (2014) (arguing that the disclosure model "represents an overly sanguine view of legislative decisionmaking").

[389] Marc Mauer, *Racial Impact Statements as a Means of Reducing Unwarranted Sentencing Disparities*, 5 OHIO ST. J. CRIM. L. 19, 42–43 (2007).

[390] *Id.* at 43. This is not to say that it is impossible though. *See* Rachel E. Barkow, *Federalism and the Politics of Sentencing*, 105 COLUM. L. REV. 1276, 1289 (2005) (arguing that impact statements have made a difference in sentencing decisions).

impact statements should not claim to do more than they are capable of doing, not that they are valueless.

3. Static Assessments. A common criticism of NEPA since its inception is that the assessment—either EIS or mitigated FONSI—is performed before project implementation and there is no monitoring to ensure the choices made work in reality.[391] This is a problem because predictions can be wrong or the facts on the ground can change, and what was once beneficial might become harmful, or a constraint imposed on the project might no longer be necessary.[392] This is a realistic issue for AISs as well, and perhaps an even bigger one. A police department is likely to update their algorithms over time. Crime statistics and demographics in a given jurisdiction will likely change over time, and demographics can change faster than environmental attributes. Having to redo an AIS might prevent necessary changes to policy, or conversely, not requiring an AIS update may allow the system to be out of alignment and discriminate more than originally predicted.

In the environmental context, many solutions have been proposed. David Hodas proposed a tort regime to place the liability for future harm on the inaccurate predictor, coupled with insurance to price it in initially.[393] Bradley Karkkainen proposed mandatory monitoring and disclosure of the results, followed by updated project plans when necessary.[394] Karakkainen's proposal is an example of "adaptive management," which has become a key concept for NEPA reformers in the environmental space.[395] Perhaps benchmarks could be created for future goals, backed up by automatically triggered defunding mechanisms if they are not

---

[391] *See* Hodas, *supra* note 351, at 188 ("NEPA's fundamental flaw is the little-appreciated fact that no one is responsible for substantive errors in EIS evaluations."); Karkkainen, *supra* note 301, at 938.

[392] *See* Karkkainen, *supra* note 296, at 938 ("[N]othing in NEPA requires or encourages the agency to engage in any follow up effort to verify the predictions made in an EA or EIS, or to adjust its decisions in light of what it subsequently learns." (footnote omitted)).

[393] Hodas, *supra* note 351, at 188–92.

[394] Karkkainen, *supra* note 296, at 938.

[395] *See* Eric Biber, *Adaptive Management and the Future of Environmental Law*, 46 AKRON L. REV. 933, 935 (2013) ("Adaptive management has become a dominant theme in the scholarship and practice of environmental law, so dominant that many scholars and managers assert that the *only* feasible option for environmental law is adaptive management."). *But see id.* at 940 (questioning the centrality of adaptive management to the future of environmental law because of its limits in reducing uncertainty and improving management and regulatory outcomes, as well as costs associated with its use).

reached. All these ideas are worth exploring for the AIS regulation. Ultimately, the broad strokes sketched here would remain the same with additional pieces to be considered later, and it is beyond the scope of this Article to evaluate those different proposals.

4. Trade Secrets. Every algorithmic accountability proposal (accountability proposals for any technology, really) eventually meets the question of how to handle the trade secret problem.[396] In short, many companies are protecting their algorithms by claiming that they are trade secrets and, therefore, cannot be disclosed.[397] Despite often being of questionable legal merit,[398] such claims are being treated credulously by courts[399] and are given great weight in public debates about "black box" technologies.[400]

---

[396] *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. (forthcoming 2018) (manuscript at 5-6) (available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2920883) ("Developers often assert that details about how their tools function are trade secrets. As a result, they claim entitlements to withhold that information from criminal defendants and their attorneys . . . ."); *see also* David S. Levine, *Secrecy and Unaccountability: Trade Secrets in Our Public Infrastructure*, 59 FLA. L. REV. 135, 139 (2007) (discussing how the private sector asserting trade secret protection when providing public infrastructure "frustrate[s] the goals of public transparency and accountability").

[397] *Cf., e.g.*, Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 5 (2014) ("No one can challenge the process of [credit] scoring and their results because the algorithms are zealously guarded trade secrets.").

[398] *See, e.g.*, Levine, *supra* note 396, at 140 ("Trade secrecy law and practices serve many useful and important purposes in private industry, but . . . their use in the public infrastructure context is inappropriate, unexpectedly powerful, and doctrinally unsound."). There is an ongoing debate about the nature of the trade secret right and whether it is derived from tort or property. *See* Mark A. Lemley, *The Surprising Virtues of Treating Trade Secrets As IP Rights*, 61 STAN. L. REV. 311, 317–27 (2008) (discussing "the strengths and weaknesses of each existing theory of trade secret protection"). Until recently, trade secrets only legally existed within a trade secret lawsuit, which includes a claim of misappropriation of said secret. *See id.* at 317. The idea of a freestanding property right on which a legally mandated disclosure would infringe is a controversial one.

[399] *See* Wexler, *supra* note 396, at 7–9.

[400] *See, e.g.*, FRANK PASQUALE, THE BLACK BOX SOCIETY 193 (2015) ("Transparency [of algorithms] was replaced by ironclad secrecy, both real and legal . . . effectively creat[ing] a property right in an algorithm without requiring its disclosure."); Frank Pasquale, *Restoring Transparency in Automated Authority*, 9 J. TELECOMM. & HIGH TECH. L. 235, 237 (2011) (discussing how creators of "black box" technologies, such as search engines or credit rating tools, circumvent the required disclosure of the patent system by asserting trade secret protection); Brenda Reddix-Smalls, *Credit Scoring and Trade Secrecy: An Algorithmic Quagmire or How the Lack of Transparency in Complex Financial Models Scuttled the Finance Market*, 12 U.C. DAVIS BUS. L.J. 87, 117 (2011) ("The principal reason for the

But none of that is altogether important to this proposal. AISs can be created without disclosing trade secrets. It will be possible to explain the rationales and choices made by policing agencies and algorithm designers without disclosing the algorithm itself. It is also possible to audit the different versions for disparate impact without disclosing the underlying algorithms by treating them as black boxes during the test.[401] Or perhaps the audits can be run—or the AISs written—by trusted public or private third parties dedicated to such tasks.[402] There are therefore ways to create the AIS without disclosing trade secrets.

Even if that were not the case, there is no reason, as a matter of policy, why trade secrets should have preferential status over something as important as fairness in criminal justice. There is no constitutional right to trade secrets,[403] and a state that wanted to pass an AIS statute could simply carve out an exception to their application. Alternatively, the AIS statute could require, as a part of procurement law, that the software must be available for audit. In sum, though there is no space here for a comprehensive discussion of trade secrets, it may appear to be a bigger problem than it actually is.

---

controversy over credit scores is that the method of calculation is kept secret from consumers, advocates, and regulators. This calculation process has been called a 'black box.' "); Sandra Wachter et al., *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT'L DATA PRIVACY L. 76, 86–88 (2017) (noting how trade secrets have limited transparency requirements in Europe's data protection law).

[401] *See e.g.*, Kroll et al., *supra* note 320, at 660 ("Beyond transparency, auditing is another strategy for verifying how a computer system works. An audit treats the decision process as a black box whose inputs and outputs are visible but whose inner workings are unseen."); *see also* Christian Sandvig et al., Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms (May 22, 2014), http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf (describing techniques for black-box audits of algorithms).

[402] *See, e.g.*, Frank Pasquale, *Beyond Innovation And Competition: The Need For Qualified Transparency in Internet Intermediaries*, 104 NW. U. L. REV. 105, 164 (2010) ("When ranking systems are highly complex and innovation is necessary . . . a dedicated governmental entity should be privy to their development and should serve as an arbiter capable of providing guidance to courts that would otherwise be unable to assess complaints about the results the algorithm generates.").

[403] *See* Ruckelshaus v. Monsanto Co., 467 U.S. 986, 1001 (1984). While *Ruckelshaus* held that trade secrets counted as property for the purpose of the Takings Clause, *id.* at 1014, the Court also noted that property rights were created by the states, not the federal Constitution. *Id.* at 1001.

5. Commodifying Discrimination. Implicit in AISs is a recognition that the discriminatory harms cannot be eliminated entirely unless the proposed course of action is blocked. As a result, the harmful impact must be weighed against the benefits it would achieve. In most contexts, that is not how people think about discrimination. Traditionally, if discrimination existed, it was to be stamped out, and inequality that need not legally be stamped out was not considered discrimination.[404] By recognizing that a certain level of discrimination might be tolerable as a tradeoff for other policing goals, there is an extent to which this proposal treats discrimination as a problem subject to cost-benefit analysis. This leads to the objection that discrimination should not be so commodified.[405]

This objection has a few distinct flavors: a moral one, a practical one, and a distributional one.[406] The moral objection claims that certain values are simply incommensurable and that engaging in the act of horse-trading on them is immoral. The practical objection notes that when amorphous normative concerns that are not easily quantified enter the realm of cost-benefit analysis, they tend to lose importance in the overall decision-making process, and that a discourse that equates these values with other goods will inevitably be willing to sell them off. Finally, the distributional objection notes that costs and benefits will be seen differently by different populations, and the powerful decision-making populations are less likely to bear a greater brunt of the cost.[407]

---

[404] *See* Michael Selmi, *Statistical Inequality and Intentional (Not Implicit) Discrimination*, 79 LAW & CONTEMP. PROBS. 199, 200–01 (2016) ("[I]n the United States, as a matter of policy, we are committed to remedying discrimination, not inequality. In other words, we will only address inequality that is the product of discrimination.").

[405] *See, e.g.*, Steven Kelman, *Cost-Benefit Analysis: An Ethical Critique*, 5 REGULATION, Jan.–Feb. 1981, at 33, 36 ("The notion of human rights involves the idea that people may make certain claims to be allowed to act in certain ways or to be treated in certain ways, even if the sum of benefits achieved thereby does not outweigh the sum of costs.").

[406] *See* Robert H. Frank, *Why is Cost-Benefit Analysis So Controversial?*, 29 J. LEGAL STUD. 913, 913–20 (2000) (discussing objections to cost-benefit analysis).

[407] There are other objections to pure cost-benefit analysis, such as the inherent uncertainty of it due to the need to inject value judgments at one time or another, rendering it arguably no better as a mode of analysis than comparing values in the abstract. *See, e.g.*, Thomas C. Heller, *The Importance of Normative Decision-Making: The Limitations of Legal Economics as a Basis for a Liberal Jurisprudence—As Illustrated by the Regulation of Vacation Home Development*, 1976 WIS. L. REV. 385, 386 (contending that regulations will often fail to achieve their objections if based only on cost-benefit analysis because many

All of these concerns are well-founded and much debated,[408] and they cannot be resolved here. The necessity of balancing incommensurable values is uncomfortable.

But necessity it is.[409] Ultimately, balancing tests are ubiquitous in the law and basically always compare two incommensurable values.[410] Cost-benefit analysis is simply a decision-making procedure that requires care in operation. Performed carefully, it will generate the same debates about how discrimination concerns should be valued as in any other discussion based in normative concerns.[411] To suggest that discrimination cannot be subject to measurement is to suggest it is a problem that can only be eradicated but not reduced. But no known decision-making process can make that happen.

As Barocas and I observed in our prior work, data mining technology forces us to reconcile with tough questions about fairness that have previously been avoided:

> [T]he pressing challenge does not lie with ensuring procedural fairness through a more thorough stamping out of prejudice and bias but rather with developing ways of reasoning to adjudicate when and what amount of disparate impact is tolerable. Abandoning a belief in the efficacy of procedural solutions leaves policy makers in an awkward position because there is no definite or consensus answer to questions about the fairness of specific outcomes. These need to be worked out on the basis of different normative principles. At some point, society will be forced to acknowledge that

---

critical variables are "economically indeterminant"). Here, I am concerned with the specific objections against turning an issue seen as a normative one into an issue of cost.

[408] *See generally* Frank, *supra* note 406, at 913.

[409] *See* John Bronsteen et al., *Well-Being Analysis vs. Cost-Benefit Analysis*, 62 DUKE L.J. 1603, 1607 (2013) ("A primary reason for [the] survival [of cost-benefit analysis] is evident and voiced often: no comparably rigorous, quantitative, and workable alternative exists for commensurating a law's positive and negative consequences.").

[410] *See, e.g.*, Patrick M. McFadden, *The Balancing Test*, 29 B.C. L. REV. 585, 624–25 (1988).

[411] *See* Matthew D. Adler & Eric A. Posner, *Rethinking Cost-Benefit Analysis*, 109 YALE L.J. 165, 167 (1999) ("CBA is a decision procedure, not a moral standard.").

> this is really a discussion about what constitutes a tolerable level of disparate impact . . . .[412]

The only real alternative to balancing is a burden-shifting framework akin to Title VII disparate impact analysis.[413] While that method avoids commodification of discrimination, it all but guarantees that no discrimination harm will outweigh a legitimate business decision. This is because, in the absence of balancing, all that is asked is whether the decision was legitimate, not if is it outweighed by other concerns. Thus, its direct impact on discrimination is worse. At least with cost-benefit analysis, there is an argument to be made that the decision is just not important enough when compared to the harm. Rejecting balancing for fear of quantifying discrimination is an example of letting the perfect be the enemy of the good.

The analogy to environmental law is helpful for considering this objection in reverse. While discrimination is seen as a distinctly moral problem, it may be necessary to sometimes treat it as measurable. Environmental concerns come from the opposite position. They have nearly always been subjected to cost-benefit analysis, yet pollution and climate change can cause massive human death tolls and should probably be considered moral issues more often than they are. There are indeed rhetorical dangers in classifying either problem as a commodity, but the language of cost-benefit analysis at least allows for the possibility that a problem can be reduced without being wholly solved and is, in principle, no different in terms of how normative concerns are valued. To let this objection stand in the way is just to blind oneself to reality.

V. CONCLUSION

Predictive policing is rapidly being adopted throughout the country, though it is unclear as of yet whether the technologies even offer any tangible benefit over traditional policing, and there is precious little insight into its discriminatory effects. Machine learning poses new regulatory challenges in many parts of society,

---

412 Barocas & Selbst, *supra* note 19, at 728.

413 *See supra* notes 260–64 and accompanying text.

but when it comes to police in particular, the track record on discrimination cries out for new forms of oversight and transparency. From "driving while black" and stop-and-frisk, to the events that led to the Black Lives Matter movement, suspicion of and violence against people of color is a consistent feature of policing in America. Police departments must ensure that they are not adopting technology that produces limited benefits while equating "criminal" with "black."[414]

The AIS provides a good starting point for regulating the disparate impact in predictive policing. Like environmental problems in the 1970s, the biggest barrier to regulation is the lack of information about the specific instances of the problem. Perhaps the benefits of predictive policing will outweigh the harms, or perhaps the harms can be mitigated. Neither advocates nor critics of predictive policing technology know the answer because the information does not exist. With AISs, police and the software manufacturers they hire will be required to produce the information that will better inform public debate, as they are the only ones that can. The AIS proposal's focus on procedural regularity and transparency allows police to take the lead and use their expertise to design efficient crime-prevention systems while requiring that they consider the externalities of their chosen course of action.

If they remain unregulated, predictive policing systems will harden and perpetuate the racial discrimination that pervades the criminal justice system. Unless society recognizes the urgency and acts soon, we will become inured to the toxic discriminatory emissions of predictive policing systems. The narrative pull of "trusting the data" will hardcode racial discrimination into the technology, making it even harder to eradicate later. Given the history of discriminatory policing, no technology or police practice should ever be adopted without investigating how it impacts minority populations. Society cannot afford to let the allure of new technologies blind people to the systemic inequalities they can perpetuate.

Impact statements are growing in popularity as a response to new and complicated technologies. In Santa Clara, California, the

---

[414] *See* Butler, *supra* note 196, at 253–54.

law requiring police to issue Anticipated Surveillance Impact Reports was passed in 2016.[415] Similar bills are pending in Oakland and New York City.[416] The European laws requiring impact assessments are also quite new. The UK's equality duty stems from a 2010 law, while the DPIAs came about in the EU's GDPR and Policing Directive that passed in 2016.[417] Oversight of predictive policing should emulate and strengthen those efforts by drawing on the original environmental regulations that spawned all of the rest. Today's data-driven technologies are simply too complicated and too important to implement without understanding the consequences for society.

---

[415] SANTA CLARA COUNTY, CAL. CODE div. A40 (2016).

[416] N.Y.C., N.Y., INT. NO. 1482-2017 (Proposed Official Draft 2017); OAKLAND CAL., SURVEILLANCE AND COMMUNITY SAFETY ORDINANCE (Proposed Official Draft 2017).

[417] Equality Act, 2010, c. 15; sch. 19; GDPR, *supra* note 307.