ARTIFICIAL INTELLIGENCE AND POLICING: HINTS IN THE *CARPENTER* DECISION

DRAFT

_ Ohio St. J. Crim. L. __ (2018)

Elizabeth E. Joh[1]

Carpenter v. United States[2] is a case riddled with ironies. A man accused of participating in the robberies of cellphone stores finds himself incriminated by his own cellphone. Several of the Justices—three of whom were born during the Great Depression—take note of the ubiquity of cellphones in our daily lives. The various opinions refer to the Cyber Age, eighteenth century dictionaries, and everything in between. For court watchers, this was a hotly debated case about the third-party exception to the Fourth Amendment that ultimately had little to do with that doctrine at all. The decision to find Fourth Amendment protections in the government's collection of a person's movements acknowledges that rapid advances in technology are changing Fourth Amendment boundaries. Lower courts are already grappling with how to apply *Carpenter*'s new protections.[3]

But let's turn instead to a different aspect of the *Carpenter* decision. On its own terms, the majority opinion resolved a "narrow" question about the government's collection of cellphone location information collected and

[1] Professor of Law, U.C. Davis School of Law. Many thanks to the Ohio State Journal of Criminal Law for the invitation to contribute to this volume.

[2] 138 S.Ct. 2206 (2018).

[3] *See, e.g.,* Naperville Smart Meter Awareness v. Naperville, __ F.3d __ (2018), available at: http://media.ca7.uscourts.gov/cgi-bin/rssExec.pl?Submit=Display&Path=Y2018/D08-16/C:16-3766:J:Kanne:aut:T:fnOp:N:2203659:S:0 (finding public utility readings at 15 minute intervals of home electricity use constitutes a "search").

Electronic copy available at: https://ssrn.com/abstract=3238212

stored by a customer's wireless carrier.[4] Chief Justice Roberts focuses on the quality of the information sought by the police as a means of deciding the case in Carpenter's favor. Less obviously, however, the majority opinion also stresses the *nature of the policing* involved in Carpenter's case: new technologies that do not just enhance human abilities. The majority makes no explicit clams about this focus. But the *Carpenter* decision reveals the Supreme Court's first set of views on how it might evaluate police use of artificial intelligence. That contention, and the questions it raises, form the subject of this essay.

TIMOTHY CARPENTER AND THE TATTLING CELLPHONE

Timothy Carpenter and his half-brother (also improbably named) Timothy Sanders assembled a loose crew of changing characters to rob seven cellphone stores near Detroit and nearby Warren, Ohio.[5] After one of the participants in the robberies confessed to the crimes and provided the cellphone numbers of the others, FBI agents applied for an order under the federal Stored Communications Act,[6] rather than a warrant premised on probable cause. The government sought cell site location information from these cellphones during the period when the crimes had occurred.[7]

That location information would have provided clues as to the defendants' whereabouts. Cellphones continuously seek a signal, usually from the closest cell site tower.[8] Each connection generates a record about the time and location of connections between a user's cellphone and a particular cell site. In that way, these records provide a detailed map of

[4] 138 S.Ct. at 2220 ("Our decision today is a narrow one.").
[5] These facts are taken from U.S. v. Carpenter, 2014 WL 943094 (E.D. Mich. 2014) and U.S. v. Carpenter, 819 F.3d 880 (6th Circ. 2016).
[6] 18 U.S.C 2703(d).
[7] 138 S.Ct. at 2212.
[8] Id. at 2211.

Electronic copy available at: https://ssrn.com/abstract=3238212

where you have been.[9] This cell site location information is stored as a matter of course by most wireless carriers.[10]

In Carpenter's case, the court orders to Metro PCS and Sprint resulted in the production of 12,898 location points cataloging Carpenter's movements during the four month period when the cellphone store robberies took place.[11] At Carpenter's trial, FBI Agent Hess created maps showing that Carpenter's phone—and him by implication—was close to the vicinity of the robberies at the time when they occurred.[12] A jury convicted Carpenter on Hobbs Act and federal firearms charges.[13]

The Sixth Circuit rejected Carpenter's claim that the collection of this cell site location information amounted to a search under the Fourth Amendment and thus required a warrant rather than a court order under the Stored Communications Act. The government's collection of this data from Carpenter's wireless carriers fell plainly, in the court's view, under the existing third-party exception to the Fourth Amendment.[14] Cell phone data—even it if revealed a time machine into Carpenter's whereabouts[15]—were no different than other business records voluntarily conveyed to third parties and thus without Fourth Amendment protection.

The Supreme Court struck a different path. The Chief Justice acknowledged the potential applicability of two lines of decisions: both the

---

[9] Id. at 2217 (noting that cell site location information provides "detailed and comprehensive record of the person's movements").

[10] Id. at 2211 (observing that wireless carriers collect this data "for their own business purposes" and that "they often sell aggregated location records to data brokers, without individual identifying information from the transmission of text messages and routine data connections.").

[11] Id. at 2212.

[12] 819 F.3d at 885 ("With the cell-site data provided by Carpenter's and Sander's wireless carriers, Hess created maps showing that Carpenter's and Sanders's phones were within a half-mile to two miles of the location of each of the robberies around the time the robberies happened.")

[13] 18 U.S.C. 924(c), 1951(a).

[14] See 819 F.3d at 888 ("This case involves business records obtained from a third party, which can only diminish the defendants' expectation of privacy in the information those records contain.").

[15] See id. at 2218 ("With access to CSLI, the Government can now travel back in time to retrace a person's whereabouts, subject only to the retention policies of the wireless carriers . . . .").

Electronic copy available at: https://ssrn.com/abstract=3238212

third-party doctrine and the privacy-in-public cases. The third-party cases like *Smith v. Maryland* emphasize that information voluntarily provided to third parties like banks or phone companies lose Fourth Amendment protection.[16] Similarly, in cases like *United States v. Knotts* the Court had denied Fourth Amendment protection for one's movements in public.[17]

Neither group of cases proved to be critical to the *Carpenter* decision. Instead, the majority opinion focused on the "unique nature of cellphone location records" to conclude that "an individual maintains a legitimate expectation of privacy in the record of his physical movements as captured through CSLI." Such information, capable of providing an "all-encompassing record of the holder's whereabouts," constitutes a "qualitatively different category" of information warranting Fourth Amendment protection.[18] The majority rebuked the government for "fail[ing] to contend with the seismic shifts in digital technology," just as it had in *Riley v. California* just four years earlier.[19]

Having concluded that the collection of cell site location data counted as a Fourth Amendment search, the *Carpenter* majority decided that a court order under the Stored Communications Act was insufficient. Absent an "urgent situation" excusing its absence, law enforcement collection of cell site location information must be obtained with a warrant.[20] By identifying a search and requiring a warrant, the decision counters the government's "powerful new tool" to investigate with a Fourth Amendment

[16] 442 U.S. 735 (1979)("When he used his phone, petitioner voluntarily conveyed numerical information to the telephone company and 'exposed' that information to its equipment in the ordinary course of business. . . . "This analysis dictates that petition can claim no legitimate expectation of privacy here.").

[17] 460 U.S. 276 (1983) ("A person traveling in an automobile on public thoroughfares has no reasonable expectation of privacy in his movements from one place to another.").

[18] 138 S.Ct. at 2216.

[19] 134 S.Ct. 2473, 2489 (2014)("The United States asserts that a search of all data stored on a cellphone is 'materially indistinguishable' from searches of these sorts of physical items. That is like saying a ride on horseback is materially indistinguishable from a flight to the moon.").

[20] Id. at 2222.

Electronic copy available at: https://ssrn.com/abstract=3238212

recalibration.[21] The majority reversed the Sixth Circuit's judgment in the case.[22]

POLICING AND ARTIFICIAL INTELLIGENCE

Plotting geographic data on a literal map for a jury evokes little of the futurism associated with the term "artificial intelligence." Loosely defined as the use of machines to approximate human thinking,[23] artificial intelligence already envelopes our daily lives—including the use of iPhone autocorrect, social media photo tagging, and the recommendations that guide what to watch, what to buy, and whom to date. The future is likely to include robotic caregivers, autonomous vehicles, and machine-driven medical diagnostics.

The availability of massive amounts of data, leaps in computing power, and increasingly sophisticated algorithms have begun to change policing as

[21] This is essentially the thesis of Orin Kerr's equilibrium adjustment theory. *See* Orin S. Kerr, *An Equilibrium-Adjustment Theory of the Fourth Amendment*, 126 Harv. L. Rev. 476 (2011). Even the author agrees: https://twitter.com/OrinKerr/status/1022573224976994304.

[22] There were four dissents in the case. Two in particular are noteworthy. Justice Kennedy found little to distinguish the collection of cell site location information from "other kinds of business records the Government has the lawful right to obtain by compulsory process." Id. at 2224 (Kennedy, J., dissenting). That analogy is remarkably similar to Kennedy's refusal to distinguish between fingerprints and DNA samples in Maryland v. King, another case in which technological advances raised questions about the limits of existing Fourth Amendment doctrine. *See* 133 S.Ct. 1958, 1976 (2013)(finding few differences between fingerprint and DNA sample collection). Justice Alito's dissent points to another important question raised in the *Carpenter* majority: whether the decision alters existing doctrines about Fourth Amendment standing. *See* id. at 2247 (noting that the Court permitting "a defendant to object to the search of a third party's property" is "revolutionary").

[23] Ryan Calo, *Artificial Intelligence Policy: A Primer and Roadmap,* 51 U.C. Davis. L. Rev. 399, 404 (2017) ("There is no straightforward, consensus definition of artificial intelligence. AI is best understood as a set of techniques aimed at approximating some aspect of human or animal cognition using machines."). Many computer scientists have understandably found fault with the imprecision with which the terms artificial intelligence and algorithms have been used in non-technical writing. For the purposes of this essay, however, the general but somewhat vague definition will have to do.

Electronic copy available at: https://ssrn.com/abstract=3238212

well. We might define the use of AI in policing as the growing use of technologies that apply algorithms to large sets of data to either assist human police work or to replace it.[24] And assistance is something of a misnomer. Artificial intelligence has begun to change the capabilities of the police, by permitting them to do what was once nearly impossible or impracticable.

One change already ushered in by artificial intelligence is an expansion in what we might call the "surveillance discretion" of the police.[25] Surveillance discretion refers to the decisional freedom of the police to pay attention to some person or persons rather than others—an uncontroversial aspect of ordinary policing.[26] Resource constraints always checked traditional surveillance discretion: there are never enough officers nor enough money for cameras and other machines. But machine-generated analyses have changed that calculus. The police today enjoy a surfeit of data that can be collected, stored, mined, and sifted through easily and cheaply: license plate data, social media posts, social networks, and soon our own faces.

The mass collection of this data would be largely useless without quick, cheap, and easy ways to find connections and patterns. Whether we call it the age of algorithms, big data, or AI, today law enforcement agencies can increasingly turn to tools that enable them to sort through this data to look for persons already identified, or for patterns from as yet unknown persons that indicate suspicious behavior. Threat analysis software might comb through private and public records to help an officer assess the potential dangerousness of a driver in a routine traffic stop.[27] Social network

[24] Elizabeth E. Joh, __ Seattle Univ. L. Rev. __ (2018).
[25] Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 Harvard L. & Pol'y Review 15 (2016).
[26] Id. at 15.
[27] *See, e.g.,* Justin Jouvenal, *The new way police are surveilling you: Calculating your threat 'score,'* WASH. POST, Jan. 10, 2016, at http://wapo.st/1OcTX3K?tid=ss_tw-bottom&utm_term=.d4d455f45465 (Intrado's Beware software provides police with a threat "score" of a person. Exactly how the software determines this is protected by Intrado as a trade secret.)

Electronic copy available at: https://ssrn.com/abstract=3238212

analysis might identify what persons pose the most likely threat of gun violence, either as perpetrator or victim.[28]

For now, we might best think of these technologies as tools that enhance the abilities of traditional policing. Predictive policing algorithms help departments focus limited human patrol resources, for instance, by highlighting people most likely to commit crimes or places where crime is mostly likely to occur.[29] But these applications can go well beyond mere enhancement. No single officer (nor a single department) can scan thousands of private and public records to make an assessment of a suspect's dangerousness.[30] No single law enforcement agency has the means to personally track every car in town and plot out its movements. No police department can deploy personnel to identify every jaywalker and fine them within seconds.[31] In these ways the tools of artificial intelligence are changing the nature of policing itself.

Another way to think of this development is that policing is becoming increasingly *automated*.[32] Automation may be most frequently associated with jobs like truck drivers, cashiers, and file clerks, but many fields will be

[28] The Chicago Police Department employs an algorithm to generate its "Strategic Subjects List." Jeff Asher & Rob Arthur, *Inside the Algorithm That Tries to Predict Gun Violence in Chicago*, N.Y. TIMES, June 13, 2017, at https://nyti.ms/2tgi63U.

[29] *See, e.g.,* John Eligon and Timothy Williams, *Police Program Aims to Pinpoint Those Most Likely to Commit Crimes,* N.Y. TIMES, Sept. 24, 2015, at https://nyti.ms/1R48saA (describing "an experiment taking place in dozens of police departments across the country, one in which the authorities have turned to complex computer algorithms to try to pinpoint the people most likely to be involved in future violent crimes — as either predator or prey"); Erica Goode, *Sending the Police Before There's a Crime,* N.Y. TIMES, Aug. 114, 2011, at https://www.nytimes.com/2011/08/16/us/16police.html (describing one program that "generates projections about which areas and windows of time are at highest risk for future crimes by analyzing and detecting patterns in years of past crime data").

[30] About half of American law enforcement agencies employ few than ten full time officers. *See* Brian Reaves, Census of State and Local Law Enforcement Agencies, 2008, 1 (2011), available at https://www.bjs.gov/content/pub/pdf/cslle08.pdf.

[31] But artificial intelligence can. *See* Christina Zhao, *Jaywalking in China: Facial Recognition Surveillance Will Soon Fine Citizens Via Text Message*, NEWSWEEK, Mar. 27, 2018.

[32] Elizabeth E. Joh, *Automated Policing*, 15 Ohio St. J. Crim. L. __ (2018).

Electronic copy available at: https://ssrn.com/abstract=3238212

subjected to varying degrees of automation.[33] This includes conventional policing. Certainly many of the most mundane tasks of patrol, including traffic direction and report writing, will be delegated to machines.[34] But even today the increasing interest in social network analysis, locational predictive policing, and threat analysis means that even those the task of assessing suspicious behavior is subject to automation as well.

ARTIFICIAL INTELLIGENCE IN *CARPENTER*

A central concern in Fourth Amendment law focuses on how the government *accesses* information. The police generally need a warrant to enter your house, whether they want to seize your most personal documents, or merely to look around.[35] In a pre-digital world, the conceptual premise of the house, mailbox, and foot locker made sense.

Finding Fourth Amendment protections for Carpenter was difficult under the Court's previous decisions because the government accessed none of the defendant's spaces normally protected by the Fourth Amendment. Instead, the majority opinion avoids this difficulty by focusing instead on the *nature* of the information sought: "the *qualitatively* different category of cell-site records."[36] Location information—at least some amount of it—can be so revealing that its very existence requires traditional Fourth Amendment protections. Much of the commentary after *Carpenter* will likely take up the question of what other information also falls into the same qualitative category of data as cellphone locational

[33] *See, e.g.*, Natalie Kitroeff, Robots could replace 1.7 million American truckers in the next decade, L.A. TIMES. Sept. 25, 2016, at http://www.latimes.com/projects/la-fi-automated-trucks-labor-20160924/ ("Trucks without human hands at the wheel could be on American roads within a decade, say analysts and industry executives.").
[34] *See* Joh, *Automated Policing*, at 2 (proposing one scenario).
[35] Orin S. Kerr, Digital Evidence and the New Criminal Procedure, 105 Colum. L. Rev. 279, 297 (2005).
[36] 138 S. Ct. at 2216 (emphasis added).

Electronic copy available at: https://ssrn.com/abstract=3238212

information and not the unprotected data that the majority loosely defines.[37]

But *Carpenter* does something yet more. The decision hints that Fourth Amendment protections also turn on the nature of the *policing* that produces the information at issue. What distinguishes the kind of policing in *Carpenter* from traditional methods also happens to describe the emerging ways in which police are relying upon artificial intelligence. *Carpenter* recognizes, perhaps more so than any other Supreme Court decision, that dramatic technological changes will rewrite the Fourth Amendment's constraints on the government's powers. In finding that we possess Fourth Amendment protections in locational data even when recorded by third parties, the Court chose to describe the data collection technique in *Carpenter* as superhuman, passive, and automated. This is noteworthy: these descriptions also characterize the very technologies of artificial intelligence that are becoming more commonplace in policing.

First, the new technologies of policing employ data collection, storage, and analysis methods that are both *superhuman and cheap*. They are superhuman because while human beings could do the same thing, it would be impracticable to do so.[38] The collection of cell site location information surpasses "the nosy neighbor who keeps an eye on comings and goings."[39] Instead, the technology is "ever alert, and [its] memory is nearly infallible."[40] Practical restraints like police staffing become much less important when there exist "tireless and absolute surveillance" methods available through technology.[41]

And even if vastly more efficient and superior to the human resources of an average police department, equally important is the affordability of

[37] Id. at 2220 ("We do not disturb the application of *Smith* and *Miller* or call into question conventional surveillance techniques and tools, such as security cameras.").
[38] Cf. "Prior to the digital age, law enforcement might have pursued a suspect for a brief stretch, but doing so 'for any extended period of time was difficult and costly and therefore rarely undertaken.'" Id. at 2217 (quoting United States v. Jones, 132 S. Ct. 945 (2012)(Alito, J., concurring in the judgment)).
[39] 138 S.Ct. at 2219.
[40] Id.
[41] Id. at 2218.

Electronic copy available at: https://ssrn.com/abstract=3238212

these technologies. Thousands of data points were available to the FBI in Carpenter's case "at practically no expense."[42] The average American police department may not possess the means to create a real time crime center but increasingly it can buy off-the-shelf software or take advantage of data already being collected by third parties. Access to these technologies is no longer an option only for the most well off municipal departments. As the Court observed in *Carpenter,* the collection of cell site location information is "remarkably *easy, cheap,* and efficient compared to traditional investigative tools."[43]

Second, artificial intelligence applications permit the expanding uses of surveillance discretion with little additional effort required from the police. With vast amounts of data being collected all the time, "police need not even know in advance whether they want to follow a particular individual or when."[44] While police will continue to seek known persons suspected of criminal activity, they will also employ "collect all" data methods to see if suspicious persons and activities "emerge" from the data. In Timothy Carpenter's case, the government was able to "access each carrier's deep repository of historical location information" "[w]ith just the click of a button."[45] These passive forms of investigation vastly expand policing power.[46]

Third, these technologies represent a decreasing emphasis on human skill in favor of automation.[47] What we might have presumed to be quintessentially human talents in policing—identifying suspicious persons and activities and drawing inferences from seemingly disconnected data—are increasingly tasks assumed by machines. No one in Carpenter's situation (i.e. anyone with a cellphone) could flee the "inescapable and

[42] Id. at 2218.
[43] Id. at 2218 (emphasis added).
[44] Id. at 2218.
[45] Id. at 2218.
[46] *See* Sarah Brayne's excellent discussion of how these technologies have change policing within the LAPD. Sarah Brayne, *Big Data Surveillance*, 82 AM. SOC. REV. 1, 14 (2017) ("The shift from query-based to alert-based systems, represents, in part . . . a fundamental transformation in surveillance activities.").
[47] Elizabeth E. Joh, Automated Policing, 15 Ohio St. J. Crim. L. __ (2018).

Electronic copy available at: https://ssrn.com/abstract=3238212

automatic nature of its collection."[48] And those tasks can be assumed at a scale, with a speed, and with results that humans could not easily reproduce. In this way, the reasoning of the *Carpenter* decision appears to recognize that police use of artificial intelligence has far surpassed merely "augmenting the sensory faculties bestowed upon them at birth."[49]

The *Carpenter* decision hints at kinds of police technologies that may necessitate new ways of thinking about the Fourth Amendment. The majority opinion does not identify these new ways of thinking definitively, but instead raises new and provocative questions. Perhaps the most direct question raised by the case will be which new technologies will qualify as "conventional surveillance techniques and tools" and thus trigger no Fourth Amendment protections.[50] The likely questions to be raised here will include the use of facial recognition technology in public spaces, particularly if they become incorporated into police body cameras intended for ordinary patrol use.[51]

The Court's decision, however, to focus not only on the quality of the information collected but also the method of policing used to obtain it suggests that novel forms of technology-enhanced policing may trigger new Fourth Amendment protections. Even if the early twenty-first century paradigm of policing is a security camera, the next will be altogether different. Many objects besides cellphones are or will be connected to the

[48] Id. at 2223 (emphasis added).
[49] United States v. Knotts, 460 U.S. 276, 282 (1983).
[50] 138 S.Ct. at 2220.
[51] The CEO of Axon, the company responsible for selling most of the police body cameras in the United States, has predicted that their cameras will soon incorporate facial recognition technology. *See* Drew Harwell, Facial recognition may be coming to a police body camera near you, WASH. POST, Apr. 26, 2018, at https://www.washingtonpost.com/news/the-switch/wp/2018/04/26/facial-recognition-may-be-coming-to-a-police-body-camera-near-you/?utm_term=.2cdaff767510 (reporting that facial recognition technology "is under active consideration" at Axon). In 2017, Axon acquired AI startup Dextro to help automate body camera video analysis. *See* Alfred Ng, *Police hear a pitch for free body cameras, with a side of AI*, CNET, Apr. 5, 2017, at https://www.cnet.com/news/police-free-body-cameras-artificial-intelligence-taser-axon-vievu/.

Electronic copy available at: https://ssrn.com/abstract=3238212

internet and each other.[52] Cloud computing will shift our perceptions of what a single source of data collection, storage, and analysis is. And increasingly our definition of policing may include responses to automated alerts. Or perhaps even automated responses to those automated alerts.

In other words, if part of what provides a person Fourth Amendment protection from surveillance is the fact that the policing method involved could be characterized as superhuman, passive, and automated, what other techniques might fall in that category? If police use 24 hour a day patrol robots capable of identifying people and vehicles in public spaces, are all of them engaged in perpetual *Carpenter*-type searches? If such robots are connected by cloud computing, how many *Carpenter*-type searches are taking place at once?

Moreover, the Court's concerns about a type of tireless, automated, and inescapable data collection would seemingly characterize "smart" cities planned for the future.[53] These visions of urban life in the future imagine an infrastructure characterized by a network of sensors intended to regulate traffic flow, respond to emergencies, and manage energy consumption.[54] Those very same sensors are also ideal methods of data collection for law enforcement as well.

Finally, if I am right about the Court's forward-looking approach to the Fourth Amendment and policing methods, that may begin to cast doubt on the extreme deference courts have given to the judgments of human police officers. Others have written extensively about the judicial reluctance to second guess police determinations of suspicion and the use of force.[55]

---

[52] *See* Andrew Ferguson, *The Internet of Things and the Fourth Amendment of Effects,* 104 CAL. L. REV., 805, 813 (2016) ("Experts predict that the worldwide scale of such 'smart' interconnected objects will continue to grow, reaching more than fifty billion objects in 2020, and one trillion by 2025. . . . "The result will be additional options for government surveillance that can reveal the patterns of everyday life.").

[53] Elizabeth E. Joh, *Policing the Smart City*, __ INT'L J. L. IN CONTEXT __ (forthcoming 2018), available at https://ssrn.com/abstract=3189089.

[54] Id. at 1.

[55] *See, e.g.,* Seth W. Stoughton, *Policing Facts,* 88 TULANE L. REV. 847, 864 (2014) ("In the context of determining whether a police use of force was constitutionally permissible, the Court has concluded that the circumstances in which police use force justify deference to the officers' decisions.").

Electronic copy available at: https://ssrn.com/abstract=3238212

Such deference presumes an accumulation of individual and institutional skill that is human, and to some extent, unknowable. But if the future of policing is automated, those assumptions may not bear their weight.

CONCLUSION

The use of artificial intelligence is nowhere to be found in the *Carpenter* decision. Indeed, the Court ends its decision eager to cabin it, "to ensure that we do not 'embarrass the future.'"[56] But in its choices to describe why the locational data obtained by the government warranted Fourth Amendment protection, the Court recognized not only the qualitatively distinct features of the information, but also the type of policing involved in obtaining it. This way of describing investigation—superhuman, passive, and automated—also happen to characterize the use of artificial intelligence in policing. And as those technologies become more powerful and prevalent in ordinary law enforcement, *Carpenter* may provide important clues regarding the Court's concerns in the future.

[56] 138 S. Ct. at 2220 (quoting Northwest Airlines, Inc. v. Minnesota, 322 U.S. 292, 300 (1944).

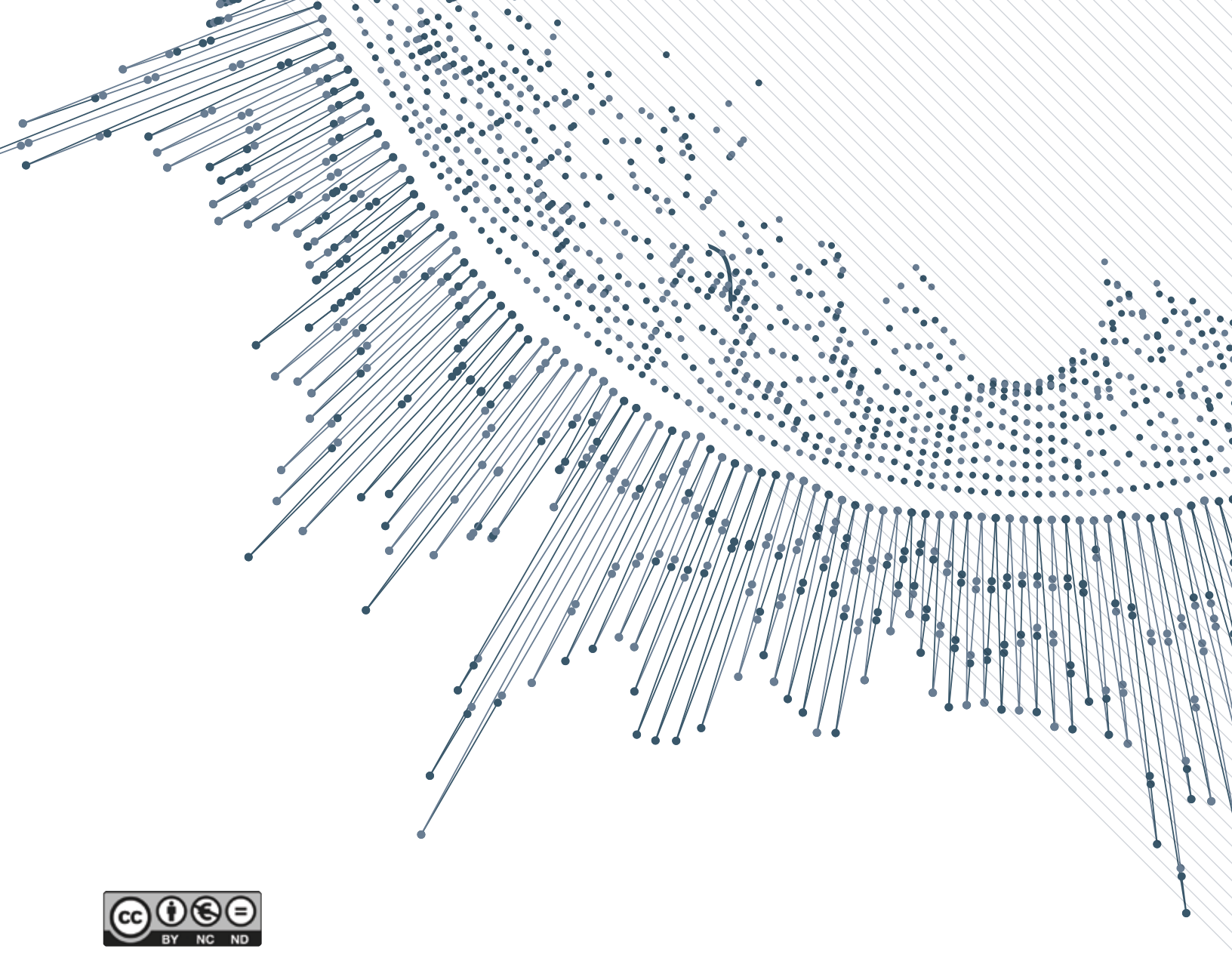
Electronic copy available at: https://ssrn.com/abstract=3238212

GARBAGE IN, GOSPEL OUT

How Data-Driven Policing Technologies Entrench Historic Racism and 'Tech-wash' Bias in the Criminal Legal System

NACDL'S TASK FORCE ON PREDICTIVE POLICING

SEPTEMBER 2021



FOR MORE INFORMATION CONTACT:

National Association of Criminal Defense Lawyers®
1660 L Street NW, 12th Floor, Washington, DC 20036
Phone 202-872-8600
www.NACDL.org/Foundation

This publication is available online at https://www.nacdl.org/datadrivenpolicing

GARBAGE IN, GOSPEL OUT

How Data-Driven Policing Technologies Entrench Historic Racism and 'Tech-wash' Bias in the Criminal Legal System

Martín Sabelli
President, NACDL
San Francisco, CA

Lisa M. Wayne
President, NFCJ
Denver, CO

Kyle O'Dowd
Interim Executive Director, NACDL & NFCJ
Washington, DC

Jumana Musa
Director of the Fourth Amendment Center, NACDL
Washington, DC

Wendy Lee
Education & Research Associate of the Fourth Amendment Center, NACDL
Washington, DC

NACDL Task Force on Predictive Policing

Cynthia Roseberry
Chair
Washington, DC

Hanni Fakhoury
Oakland, CA

Juval Scott
Charlottesville, VA

Robert Toale
New Orleans, LA

Bill Wolf
Chicago, IL

Reporter

Michael Pinard
Professor of Law and Co-Director, Clinical Law Program
University of Maryland, Francis King Carey School of Law

Co-authors

Wendy Lee, Jumana Musa, and **Michael Pinard**

TABLE OF CONTENTS

ABOUT THE NATIONAL ASSOCIATION OF CRIMINAL DEFENSE LAWYERS AND THE NACDL FOUNDATION FOR CRIMINAL JUSTICE

The National Association of Criminal Defense Lawyers (NACDL) is the preeminent organization in the United States advancing the goal of the criminal defense bar to ensure justice and due process for persons charged with a crime or wrongdoing. NACDL envisions a society where all individuals receive fair, rational, and humane treatment within the criminal legal system.

NACDL's mission is to serve as a leader, alongside diverse coalitions, in identifying and reforming flaws and inequities in the criminal legal system, and redressing systemic racism, and ensuring that its members and others in the criminal defense bar are fully equipped to serve all accused persons at the highest level.

NACDL members — and its 90 state, local and international affiliates — include private criminal defense lawyers, public defenders, active U.S. military defense counsel, and law professors committed to promoting fairness in America's criminal legal system. Representing thousands of criminal defense attorneys who know firsthand the inadequacies of the current system, NACDL is recognized domestically and internationally for its expertise on criminal justice policies and practices.

The NACDL Foundation for Criminal Justice (NFCJ) is a 501(c)(3) charitable non-profit organized to preserve and promote the core values of the American criminal legal system guaranteed by the Constitution — among them access to effective counsel, due process, freedom from unreasonable search and seizure, the right to a jury trial, and fair sentencing. The NFCJ supports NACDL's efforts to promote its mission through resources education, training and advocacy tools for the public, the nation's criminal defense bar, and the clients they serve.

NATIONAL ASSOCIATION OF CRIMINAL DEFENSE LAWYERS | NACDL FOUNDATION FOR CRIMINAL JUSTICE

National Association of Criminal Defense Lawyers®
1660 L Street NW, 12th Floor, Washington, DC 20036
Phone 202-872-8600
www.NACDL.org/Foundation

This publication is available online at https://www.nacdl.org/datadrivenpolicing

FOREWORD

In the summer of 2020, it was estimated that as many as 26 million people took to the streets to protest not only the killing of George Floyd, but the systemic racial violence that American policing had come to represent in the eyes of millions of people around the nation and around the world. During this time, thousands called for alternatives to punitive policing and for an end to police brutality. Subsequently, the onus shifted to the mayors, the city council people, the police chiefs — the managers of the policing systems around the nation — to respond. In many instances, the temptation to elude the discussion of race and embrace the ever-present hope that technology could offer a panacea to our society's most fundamental racial justice problem proved too difficult for many of these system managers to resist.

In times of crisis, law enforcement agencies increasingly turn to technology instead of instituting meaningful reforms. Immediately, after police officers shot and killed Michael Brown in Ferguson, Missouri, law enforcement agencies were encouraged to adopt body cameras as a potential "solution" to state violence. Yet the number of people killed by law enforcement, and the disproportionate representation of Black and Brown people in those figures, has not changed in the intervening seven years. Although already heavily in use at the time, the international uprising against racism and state violence in the United States in the aftermath of the killing of George Floyd, provided an opportunity to once again under the guise of technological innovation use public outcry for reform as a tool to increase police power. These technologies use algorithms to mine vast troves of data, directing law enforcement on where and who to police with the purported intention of removing bias from policing. As a result, system managers around the nation too timid to look squarely into the face of the racial realities highlighted by the protesters have again latched on to a faux technology grounded panacea.

Today, data-driven policing, buttressed by easy access to previously unimaginable data sets, is a burgeoning movement that has been implemented by dozens of cities. It has facilitated more aggressive policing measures than ever before and as opposed to "defunding the police," it offers a "solution" that provides grounds for dedicating even more resources to police departments, as body cameras did seven years ago.

Also, with over $100 billion spent annually on policing in the United States, it further drains resources that could otherwise go towards supporting more effective methods of improving public safety, such as improving healthcare, nonviolent crisis intervention programming, or climate justice efforts that could save lives lost during increasingly common extreme weather events. And by embracing a pro-active policing model which encourages the predicting of crime before it happens, data-driven policing facilitates the reification of long-standing racial bias and allows the use of force to spike as it now will be justified by the veneer of science. Ironically, as the millions of people who took to the streets called for investment into alternatives to policing and an end to police violence, the data-driven policing response seems to represent the defiant antithesis to the will of the people.

In contrast, this report by the National Association of Criminal Defense Lawyers (NACDL) more properly responds to the moment by engaging in a rigorous, thorough analysis of data-driven policing. It ultimately

calls for the abandonment of its adoption, and transparency in jurisdictions where the train has already left the station. By reminding decision-makers of the impact that these systems might have on constitutional rights, privacy, and the ability of our youth the thrive, NACDL has done a true public service that I can only hope will be disseminated far and wide.

In 1967, President Lyndon B. Johnson set up a commission to understand a wave of protests following an incident of police brutality in Detroit that resulted in over 7,000 arrests. That 1967 report, called the Kerner report, found that police "symbolize white power, white racism, and white repression" for significant numbers of Black people.

Over 55 years later, the problems seem to have only intensified. The Department of Justice has conducted more than 70 investigations of some of the nation's largest local municipal policing regimes, finding a wide array of evidence indicating deeply held racial hostility persists. Meanwhile, the broader impact of racially targeted policing endures. Black Americans represent 13% of the U.S. population, yet comprise 40% of those incarcerated, and Black Americans are accordingly 2.5 times more likely to be arrested than their white counterparts. This disparity is particularly troublesome for crimes like drug possession, where studies confirm that 5% of illicit drug users are Black, yet they represent 29% of those arrested and 33% of those incarcerated for drug offences. In other words, the disparities in punishment are not indicative of disparities in behavior. This is what systemic racism looks like.

The nation's criminal defense lawyers, more so than perhaps any other group in the nation, have the knowledge, expertise, and power to bring this racially targeted policing regime to light and serve as a bridge between today's sad reality and the hope for our people to be able to live their lives to the best of their capacity in an environment that provides dignity, community, and a sense of justice. This is part of that effort, and I hope that you continue the work of building on the important accomplishment that this report represents.

Best Wishes,

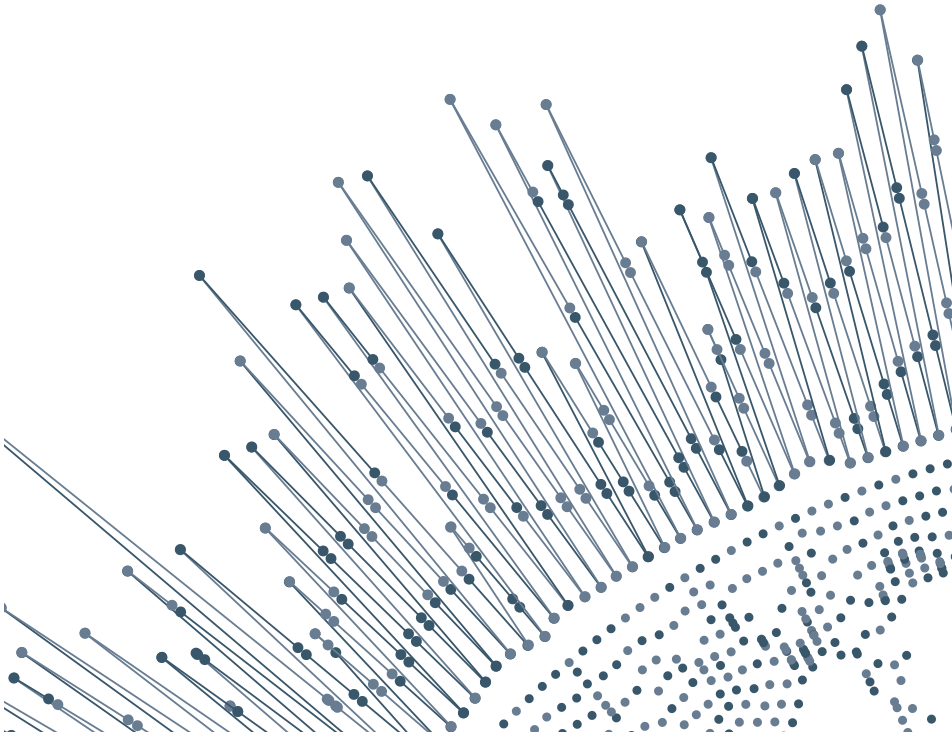
Justin Hansford
Professor, Howard University School of Law
Executive Director, Thurgood Marshall Civil Rights Center

ACKNOWLEDGMENTS

Thanks also to the many experts who made themselves available for interviews or responded to requests for information, including Shahid Buttar, P. Jeffrey Brantingham, Sarah Brayne, Matt Cagle, Henrik Chulu, Cynthia Conti-Cook, Malkia Cyril, Jarrell Daniels, Andrew Ferguson, Zach Friend, Michelle Fields, Jamie Garcia, Jeremy Heffner, Brian Hofer, Chaclyn Hunt, Daniel Kahn Gillmor, Aziz Huq, Elizabeth Joh, Hamid Khan, Nitin Kohli, Young-Mi Lee, Rachel Levinson-Waldman, Scott Levy, Kristian Lum, Sean Malinowski, Chad Marlow, Freddy Martinez, Brian McDonald, Matt Mitchell, Darleene Murray, Taylonn Murphy, Cathy O'Neil, John Patzakis, Anthony Posada, Myla Rahman, John Raphling, David Robinson, Jeffrey Ratcliffe, Steven Renderos, Rashida Richardson, Jessica Saunders, Jay Stanley, Philip Stark, Vincent Southerland, Josmar Trujillo, Stephanie Ueberall, Kevin Vogeltanz, Rebecca Wexler, and Pete White.

Thanks also to Rachel Levinson-Waldman, Rashida Richardson, Vincent Southerland, and Harlan Yu, who reviewed and consulted on this report as it was being drafted. Thank you to Kian Vesteinsson for reviewing the report and for working with the Task Force as a Former Education and Research Associate for the Fourth Amendment Center.

Grateful thanks also to the various individuals at NACDL who provided feedback and guidance, including Kyle O'Dowd, Interim Executive Director of NACDL; Norman L. Reimer, Senior Policy Consultant at NACDL; Ivan Dominguez, Senior Director of Public Affairs and Communications; Michael Price, Litigation Director for the Fourth Amendment Center; Aisha Dennis, Staff Attorney for the Fourth Amendment Center; Priyanka Podugu, Education and Research Associate for the Fourth Amendment Center, Jordan Murov-Goodman, Former Legal Fellow for the Fourth Amendment Center; and Kaitlin Galindo, Former Legal Intern for the Fourth Amendment Center.



EXECUTIVE SUMMARY

We find ourselves at a watershed moment for policing in the United States. Sparked by the continued tragic killings of unarmed Black men, women, and children by police officers – and particularly the galvanizing protests, demands, and efforts that have carried forward from the killings of George Floyd and Breonna Taylor – calls to reform, transform, defund, dismantle, and abolish law enforcement have taken center stage in cities, towns, and legislative halls throughout the United States. The families of victims who have been killed by police officers, along with other impacted residents, activists, civil rights organizations, elected officials, government agencies, and policing professionals, among many others, are pressing the need for radical change. Journalists and news organizations are reporting on and amplifying the voices for transformation.

These actions and efforts have crystalized in response to policing practices that have harmed and hyper-criminalized individuals and communities of color. Over the last several years, U.S. Department of Justice investigations have revealed what residents of Black and Brown communities have long known and expressed: that abusive and unconstitutional policing pervade their lives. The National Association of Criminal Defense Lawyers (NACDL) convened a Task Force on Predictive Policing[1] in 2017 to examine the rise, implementation, and ramifications of data-driven policing technologies, and to offer comprehensive recommendations that support transparency, safety, and due process.

From 2017 to 2019, the Task Force held meetings in Washington, D.C., Chicago, Los Angeles, San Francisco, and New York City. At these meetings, the Task Force met with a diverse array of witnesses, including technologists and industry experts, law enforcement personnel, academics, attorneys, advocates, and community stakeholders. The purpose of these meetings was to learn about the different types of technologies that police departments currently use or have used in communities throughout the country, and the impact of these technologies on individuals, families, communities, and the criminal legal system.

In recent years, police departments have been turning to and relying on rapidly developing data-driven policing technologies to surveil communities, track individuals and, purportedly, predict crime. These technologies include algorithmic decision-making that departments claim can predict where crime is likely to occur, who will likely commit crime, and who will likely be a victim. These algorithms are thus designed to interrogate massive troves of data gathered in a myriad of ways, using inputs that can range from police-generated crime reports to publicly available social media posts. The outputs are then used to make critical decisions about patrols, or to make life-altering designations of individuals.

The Task Force chose to use the term "data-driven policing" to refer to the various technologies and practices that were studied in order to avoid the shifting sands of terminology. In doing so, the Task Force understands that other technologies that do not utilize automated decision-making systems can fall under this broad terminology, but for the purpose of this report, data-driven policing is used to refer to the tools that analyze data to determine where, how, and who to police. During the time that the Task Force studied the issue, Palantir, a company that contracts with law enforcement agencies in major cities across the country, never described its tools as "predictive policing." The company instead advertises as a "data integration and analysis software platform."[2] As

the term predictive policing fell out of favor, NYPD rolled out "precision policing," which "combined predictive policing analysis and community policing."[3]

This Report identifies the racial overtones present in the historical development of data-driven policing systems and the data they utilize, and looks to case studies of specific predictive policing systems to illuminate how this technology invades personal privacy and operates as a confirmation-bias tool to justify over-policing low-income communities of color. The purpose of this Report is to: (1) call attention to the rapid development and deployment of data-driven policing; (2) situate data-driven policing within the racialized historical context of policing and the criminal legal system; (3) make actionable recommendations that respond to the reality, enormity, and impact of data-driven policing; and (4) suggest strategies for defense lawyers in places where data-driven policing technology is employed.

Brief History of Policing and the Economics of Punishment

Modern policing traces its origins to "slave patrols," a tool of racist oppression. Policing has been stitched into the fabric of the country since its founding over 400 years ago. As both a practice and a profession, policing has evolved over the centuries. Even after the Civil War and the Thirteenth Amendment, a loophole in the Amendment's text legalizing enslavement as a "punishment for crime" turned antebellum slave patrols into a policing and criminal legal system designed to criminalize, incarcerate, and otherwise punish Black people in ways that strongly resembled and replicated slavery. Black Codes and vagrancy statutes were used to hyper-criminalize, imprison, and exploit Black people for financial gain.[4]

When the criminal legal system began shifting to models rooted in incapacitation and deterrence during the 1970s, individuals and communities without social and political capital increasingly lost access to law enforcement services. Police departments were overwhelming white, were largely unresponsive, and often showed outright hostility to the needs of Black residents. In the decades that followed, policing strategies focused on zero tolerance, broken-windows, and an escalation of the War on Drugs, and systematically targeted Black communities for surveillance and incarceration. Following the 9/11 terror attacks, police departments militarized in the name of national security and terrorism prevention, but the result was heightened surveillance of the public, particularly communities of color.

One can draw a direct line from this history to the hyper-criminalization of Black men, women, and children today, along with the economic incentives embedded in the modern criminal legal system. Police officers continue to serve as arbiters of who is introduced to the system and who is not. Law enforcement strategies and tactics that criminalize and capture poor Black, Indigenous and People of Color ("BIPOC") men, women, and children, continue. A historical analysis of policing thus demonstrates two aspects of policing that are constant over time: first, policing is rooted in the control and criminalization of Black people; and second, lawmakers, prosecutors, and courts have historically given deference to law enforcement concerning their practices.

Brief History of Surveillance and the Rise of Big Data

An estimated 2.5 quintillion bytes of data are now generated every day.[5] With the ability to digitize, track, and store nearly every aspect of information, from Internet searches and social media posts to cell phone calls and retail purchases,[6] individuals are constantly contributing "to a growing trove of data as they go about their daily lives,"[7] and such data plays an ever-expanding role in the surveillance of individuals and communities, and in determining who, where, and when police officers monitor, encounter, search, and arrest.

Contemporary data-driven programs in policing emerged from repeated attempts by criminologists, legal scholars, and law enforcement agencies to quantify and measure the complex social processes behind crime and disorder. Though these programs may vary in the types of data and techniques they employ, all of them are inevitably shaped by prior policing patterns, historical crime reports, and other records compiled by the police themselves. As such, data-driven policing obscures the discretion, biases and human decision-making inherent in the production of such data.[8]

The technological tools used to organize and interrogate police data are the natural extension of tactics that have been employed over centuries of policing. Data collected and organized by law enforcement agencies have long-informed police actions, and reflect the implicit bias and explicit racism of policing as an institution. Using algorithmic tools to interrogate massive troves of police data does not correct for biased or racist policing practices; rather, it entrenches them. As such, the technological tools that are guiding modern day policing practices must be analyzed within the historical context that birthed them.

Modern-day policing is fueled by an almost unfettered access to immeasurable amounts of personal data, ensuring that officers have seamless access to criminal intelligence at the local, state, and federal level. As such, law enforcement databases are frequently exempt from complying with the same constitutional and legal standards that govern criminal investigations. Even with information that would normally require a warrant, law enforcement agencies can purchase the data from commercial data brokers without the need for a subpoena or warrant. With such an unfathomably massive amount of data being constantly generated, police departments are perpetually at risk of employing predictive algorithms trained on erroneous or irrelevant data, or even data manipulated by the police themselves. Research and scholarship have repeatedly found that these databases are often riddled with errors, and that "biases in the databases themselves, based on how data are collected, may also lead to disparate outcomes."[9] Indeed, all data-driven policing systems run the risk of being built on an incomplete and biased understanding of where crimes take place and who is actually committing them, with real costs to communities already under pervasive police scrutiny and surveillance.

The Landscape of Data-Driven Policing

Today, data-driven policing encompasses the many surveillance technologies, tools, and methods employed by police officers to visualize crime, target "at-risk" individuals and groups, map physical locations, track digital communications, and even collect data on individuals and communities. This can include any approach that incorporates a clear reliance on information technology, criminology theory, and predictive methods in policing.[10] At its core, predictive algorithms in policing programs are the "data-driven incarnation"[11] of what criminologists have been attempting to achieve for decades: to analyze past events, infer broader patterns, and to then use those insights to "prevent" future crime.

According to a report published by the RAND Corporation, predictive methods in policing can generally be divided into four broad categories: (1) methods for predicting crimes, or approaches used to forecast places and times with an increased risk of crime; (2) methods for predicting offenders, or approaches that identify individuals at risk of offending in the future; (3) methods for predicting perpetrators' identities by creating profiles that accurately match likely offenders with specific past crimes; and (4) methods for predicting victims of crimes by identifying groups, or in some cases, individuals who are most likely to become victims of crime.[12]

In order to implement these methods, predictive policing employs a variety of machine learning algorithms. Since the developers of data-driven policing technologies often assert trade secret evidentiary privileges to deny public access to the inner workings of their algorithms, the types of machine learning used in such programs are relatively unknown, and because many of these tools built on these algorithms are relatively new, or are continuing to change alongside advancements in technology, all are "are relatively untested, with only a handful of studies, reports, or empirical validation across jurisdictions."[13] Once predictions are made, "there is, generally, no standard for how police should use the predictions,"[14] meaning the technology gives an objective façade to more traditional policing tactics.

Place-based data-driven policing programs are built upon the premise that crime is not evenly dispersed geographically, and that certain places are expected to experience higher rates of crime over a certain period of time. Like "hot spot policing," or the identification of geographically-bound spaces associated with a proportionally greater number of criminal incidents or heightened victimization risk, place-based crime forecasting visualizes the spatial and temporal distribution of crime to purportedly "predict" areas with future criminal activity. NACDL found that the use of predictive algorithms in place-based crime forecasting produced harmful, self-perpetuating feedback loops of crime predictions, in which officers would repeatedly patrol neighborhoods

that had been disproportionately targeted by law enforcement in the past, and were thus overrepresented in the historical crime data used to train and build predictive crime algorithms.

Police departments also rely on person-based data-driven policing programs to predict who is most likely to commit crimes, in addition to who is likely to become the victim of a crime. The algorithms behind these programs are designed to interrogate massive troves of data gathered in a myriad of ways, with inputs ranging from police-generated crime reports to publicly available social media posts. The outputs are then used to make critical decisions about patrols, or life-altering designations about which individuals need to be suspected, surveilled, and encountered by law enforcement. These programs are enabled by law-enforcement databases and have been shown to lead to the increased enforcement and arrests of predominantly Black and Brown young men.

Among these databases are gang databases, which are localized within cities or similar jurisdictions, and encompass a broad swath of identifying data on individuals known, suspected, believed, or assumed to be gang members, associated with gang members, or affiliated with gang members. BIPOC communities comprise the vast majority of individuals listed on these databases, with Black men being the most overrepresented group. Individuals can be certified as gang members simply based on their appearance or location, or even their likes, comments and connections on social media, often without being notified of their inclusion in such a database or given the opportunity to challenge that designation. In addition to being over-inclusive, hyper-racialized, and non-transparent, NACDL found that these databases are riddled with errors, and have even included young children and infants.

Though data-driven policing programs may be designed to lower citywide violence levels and marketed as intervention opportunities for the benefit of communities, empirical research studies have repeatedly found that such tools fundamentally remain a law enforcement deterrence tool. Some programs have resulted in heightened risk of arrest, in addition to enhanced federal and state sentencing options, for designated individuals sweapt into their broad net. For example, individuals included in gang databases are subject to increased police surveillance and monitoring, and can also face enhanced criminal charges upon arrest.

Person-based data-driven policing programs have historically been shrouded in secrecy, with police departments frequently using social media monitoring tools and techniques to surveil individuals, groups, and communities without their knowledge. Data obtained through social media posts and text messages are increasingly being used to not only populate gang databases, but as primary evidence in criminal investigations, with no effective means of oversight to limit the extent of surveillance.

Critical Analysis of Data-Driven Policing

Methodological Problems

Though prediction has always been a fundamental part of policing,[15] the emergence of predictive algorithms in policing was considered particularly novel for its alleged ability to apply artificial intelligence to quantities of data once considered too large and too complex for police departments to analyze.[16] Its proponents have since claimed that predictive policing programs can lower crime, revolutionize public safety, and help under-resourced departments "do more with less,"[17] while critics have argued that such programs produce self-perpetuating feedback loops of crime prediction, placing historically over-policed individuals and communities at further risk of harm. The Task Force found the latter, that these programs entrench existing biases and exacerbate the disproportionate impact of policing on BIPOC, low income and other marginalized communities.

If crime data is to be understood as a "by-product of police activity,"[18] then any predictive algorithms trained on this data would be predicting future policing, not future crime.[19] Neighborhoods that have been disproportionately targeted by law enforcement in the past will be overrepresented in a crime dataset, and officers will become increasingly likely to patrol these same areas in order to "observe new criminal acts that confirm their prior beliefs regarding the distributions of criminal activity." [20] As the algorithm becomes increasingly confident that these locations are most likely to experience further criminal activity,[21] the volume of arrests in these areas

will continue to rise, fueling a never-ending cycle of distorted enforcement.[22] The biases held by police officers and those reporting crimes, and correlations between attributes like race and arrest rates, will not only be recognized and replicated by the algorithm, but directly integrated into the software "in a way that is subtle, unintentional, and difficult to correct, because it is often not the result of an active choice by the programmer."[23]

With an increasing number of police departments already succumbing to the "pressures of managerial techniques that emphasize quantitative measures of effective policing,"[24] some experts have suggested that data-driven policing strategies and tools have facilitated the return of broken windows policing. Since people also have a tendency to believe a computer-generated report over that of a human-created report, predictive policing programs and other automated decision-making systems often run the risk of "being trusted above human judgment while simultaneously concealing potential unchecked errors."[25] Biases in machine learning algorithms pose a "particularly insidious risk to disadvantaged groups by creating a pseudo-scientific justification for discriminatory treatment,"[26] and while transparency can help prevent deliberate or semi-deliberate discrimination, it cannot singlehandedly "correct the effects of the unintentional, institutional discrimination embedded in the data itself."[27]

Transparency, Trade Secrets, and Non-Disclosure Agreements

Intensified public scrutiny of these predictive algorithms has raised questions about how they are developed, implemented, and marketed; why they are not subject to more review; and whether there are mechanisms in place to properly assess their risks, vulnerabilities, and potential for greater societal harm. Moreover, the private companies that build, market, and sell data-driven policing technologies not only claim propriety rights over their methodologies, but assert such claims in response to subpoenas;[28] effectively denying defense attorneys and those accused in criminal cases information central to the defense.[29]

The lack of transparency with data-driven policing and its fundamental business model are deeply interwoven. With private companies competing to build, market, and sell technology, police departments "are customers or clients of private companies"[30] — with some companies even providing technology to police departments initially for free, with the goal of selling departments on the need to continue using the technology. As a result, many departments often possess little to no insight into the inner workings of the systems they employ and lack incentive to do so without explicit transparency measures in place. This lack of transparency jeopardizes fundamental constitutional rights, public trust, and privacy, and cedes too much control to companies in the private sector. This lack of transparency also extends to the adoption and use of the technology, and the ways in which it undermines democratic governance. Most police departments do not inform impacted communities, let alone legislators, that they are utilizing data-driven policing technologies, and rarely provide justification or disclose the policies that govern a technology's use.

While policing as a practice remains largely unchanged, the decisions that guide law enforcement today are being outsourced to private entities with no perceived obligation to publicly disclose details of how their tools actually work. It is consistent with historical trends that police departments are unchecked in their use of expanding and invasive technology to surveil the public, and that law enforcement deploys this technology in ways that continue to hyper-criminalize Black and Latinx individuals. By "tech-washing" racially biased policing practices and hiding behind data-driven tools that collect, use, and produce skewed data, law enforcement agencies are able to justify increased policing and surveillance in historically over-policed communities under the veneer of technological neutrality and objectivity. In this way, data-driven policing perpetuates a self-reinforcing cycle of bias and inequity.

Impact on Youth

Children possess special protections in the juvenile court system, such as different sentencing guidelines, an emphasis on rehabilitation over punishment, and criminal records that are sealed and typically expunged once they turn eighteen years of age. In spite of this, many continue to be criminalized by highly secretive data-driven policing technologies, tools, and programs that cause lifelong collateral consequences. These inscrutable systems have been documented to be racially skewed, are riddled with errors, and have historically

included children as young as eleven years old. Moreover, users rarely notify minors of their inclusion or offer the ability to seek their removal from such systems.

Since the inception of these databases, "police officers have been racially profiling and tracking people – primarily youth of color – suspected of 'gang involvement' often based on what they look like, where they live, and how they dress."[31] These databases also allow law enforcement officers to share extensive information about gangs, and to "collect, store, and analyze personal information about alleged gang members;"[32] with many of them "filled with the names and pictures of thousands of young people of color who have not been convicted of any crimes."[33] As a result of these data-driven policing technologies, tools, and programs, many children continue to be treated as adults in the criminal legal system, in violation of their fundamental rights to special protection and to be tried by a specialized juvenile justice system.

For example, CalGang, a database widely used in California, listed 42 infants under the age of 1 as active gang members.[34] Moreover, because there is "no clear, consistent and transparent exit process" for those on the database, it can be assumed that a significant proportion of "gang" designees were added in their teens and preteens.[35] The Chicago Police Department (CPD)'s database includes more than 7,700 people who were added to the database before they turned 18, including 52 children who were only 11 or 12 years old at the time of their inclusion.[36] An investigation published by *The Intercept* identified hundreds of children between the ages of 13 and 16 listed in the New York Police Department (NYPD)'s gang database in 2018.[37] The Boston Police Department (BPD) uses a point system to determine whether to include someone in its "Gang Assessment Database";[38] making it possible for teenagers to be designated as gang members "simply because of the people they're being seen with,"[39] and without any actual allegation of violence or criminal activity.

This provides "disturbing insights into the police targeting of young people,"[40] and the ease with which officers can add a minor to a database for having a tattoo symbolizing a gang, for wearing clothing associated with a gang, or for repeatedly visiting "a gang area."[41] Because of the secrecy surrounding gang databases, some have even referred to them as hidden "surveillance tool[s] for monitoring children;"[42] with such monitoring often taking place on social media, where officers can search a user's publicly available account and posts; establish an undercover account to interact with a targeted user; or use a search warrant to get additional information about a specific user.[43]

Critics have additionally argued that gang databases — with opaque methods used to obtain intelligence and data for such systems and little information available on how someone gets on or off these lists — function like "black boxes," making them a prime tool for racial profiling.[44] Studies have also shown that once an individual is listed in a gang database, they will likely encounter increased police attention and harassment. Since gang databases make gang identification information significantly more accessible to law enforcement officers, this has resulted in the more widespread stopping of young people of color, even without suspicion of criminal activity.[45]

Impact on Constitutional Rights and the Criminal Process

Data-driven policing has proliferated so quickly that solutions lag for the myriad constitutional rights that are implicated by its deployment and use. The aggregation and classification of vast and disparate types of personal information raises serious concerns about the First, Fourth, Fifth, Sixth, and Fourteenth Amendment rights for those suspected and accused in criminal cases.

Data-driven policing raises serious questions for a Fourth Amendment analysis. Prior to initiating an investigative stop, law enforcement typically must have either reasonable suspicion[46] or probable cause.[47] Does a person loitering on a corner in an identified "hotspot" translate to reasonable suspicion? What if that person was identified by an algorithm as a gang member or someone likely to be involved in drug dealing or gun violence? Can an algorithm alone ever satisfy the probable cause or reasonable suspicion requirement?[48] The lack of transparency and clarity on the role that predictive algorithms play in supporting reasonable suspicion determinations could make it nearly impossible to surface a Fourth Amendment challenge while replicating historic patterns of over-policing.

Data-driven policing databases may also work an end run around the Fourth Amendment by making law enforcement privy to information that would otherwise require a warrant to access. The government rarely discloses its use of data-driven policing technologies. Even if the use of the technology is a matter of public record, the inputs used, training data, and algorithms are proprietary and therefore shielded from scrutiny. This raises a number of due process issues that implicate a person's right to a fair trial.

In *Brady v. Maryland*,[49] the Supreme Court found that the government has an obligation to provide defendants with evidence that is material to a determination of either guilt or punishment. The government's failure to disclose the use of certain technologies or databases may raise a *Brady* issue since defense attorneys will not have the opportunity to challenge whether the results or the tools themselves were inaccurate or improperly deployed.

Algorithmic tools often use claims of proprietary software and trade secrets to shield their technology from outside scrutiny. The companies that develop the tools conduct their own validation studies, rather than rely on independent verification and validation, which is the accepted practice. Allowing companies with a financial interest in the success of their tools to validate their own technologies with no outside scrutiny is scientifically suspect.[50] It also frustrates the ability of the defense to challenge the reliability of the science underlying the novel software.

As the Task Force heard throughout their investigation, data-driven policing tools often reinforce or even exacerbate the racial biases that have always existed in policing. In other words, the government's use of data-driven policing software has a disparate impact on individuals of different races. Problematically, technological tools often enhance the discriminatory effect even as they make it more difficult for individuals to bring an Equal Protection claim. According to legal scholar Aziz Huq:

> The concerns of constitutional law simply do not map onto the ways in which race impinges on algorithmic criminal justice. The result is a gap between their legal criteria and their objects.... The replacement of unstructured discretion with algorithmic precision, therefore, thoroughly destabilizes how equal protection doctrine works on the ground.[51]

Part of the issue is that an Equal Protection claim against facially neutral government action requires that a litigant show discriminatory intent as a threshold element.[52] Data-driven tools are designed in a manner whereby bias is buried beneath the technology. Because any bias is filtered through an algorithm, critics have accused data-driven tools of "techwashing"[53] the biases inherent in the data. There is an inherent conflict between the reality that machine learning is not advanced enough to formulate intent, and the fact that "the unthinking use of algorithmic instruments will reinforce historical race-based patterns of policing."[54]

Although First Amendment concerns are not primary in criminal prosecutions, there are several First Amendment issues raised by data-driven policing programs and technologies. When people are criminalized based on their associations and their participation on social media, they are subject to what Professor Elizabeth Joh calls the "surveillance tax." As Joh writes, the intrusiveness of surveillance extends beyond arrest: "Knowledge of surveillance alone can inhibit our ability to engage in free expression, movement, and unconventional behavior."[55]

Gang designations and inclusion on lists of potential offenders are often based on proximity, associations, social media interactions, comments and posts rather than facts and evidence. The low bar for inclusion in such databases, the lack of notice, the inability to challenge one's inclusion on the list and the real-world consequences of such designations create circumstances where young people are forced to live with the potentially life-changing consequences of such designations based on communications that should be protected speech under the First Amendment.

TASK FORCE RECOMMENDATIONS ON DATA-DRIVEN POLICING TECHNOLOGIES

1. Top-Line Recommendation

Police departments must not utilize data-driven policing technologies[56] because they are ineffective; lack scientific validity; create, replicate and exacerbate "self-perpetuating cycles of bias";[57] deeply entrench

existing inequities in the system; hyper-criminalize individuals, families, and communities of color; and divert resources and funds from communities that should be allocated towards social services and community-led public safety initiatives.

While the Task Force believes these technologies should never be used, it is clear that these technologies are being considered or have been implemented in cities and towns across the country. Lack of access and transparency will hamper defense lawyers' ability to properly represent their clients. The following recommendations are for areas that are already using these technologies. These recommendations are in no way intended to serve as principles for implementing such technologies. Rather, they are mitigation efforts intended to ensure the most transparency and equity for people ensnared by these technologies, and to give defense attorneys the notice and transparency they need to defend their clients.

2. Governing Use

Police departments seeking these tools must not adopt any data-driven policing technology without first meaningfully engaging the communities where it would be deployed and without first securing approval for the technology from the elected governing bodies that represent the impacted communities.

This process must include the residents of the communities where the data-driven policing technologies would be deployed, community organizations, organizations focused on youth from the impacted communities, and attorneys with expertise in upholding the constitutional rights and civil liberties of residents from impacted communities.

As part of engaging impacted communities about the proposed data-driven technology, resources must be allocated to local governing bodies to host forums to present and describe the proposed law enforcement technology to the residents of the impacted communities. These forums would also provide a space for impacted communities and law enforcement to discuss the law enforcement need for the proposed technology, detailing how the policies governing the use, scope, and limitations of the technologies would be implemented within the defined law enforcement need. Resources and space should also be allocated to enable and empower community members to provide feedback about the technology, and to address community concerns about transparency, racial bias, and the impact of the proposed technology on civil liberties and constitutional rights. If there is a majority consensus by state or local governments and impacted communities that the proposed technology should not be used by law enforcement, then the technology should be prohibited.

3. Transparency

Prior to implementing any data-driven policing technology, law enforcement must adopt written policies governing the technology's use. Before adopting these policies, law enforcement departments must make draft policies available to the public, provide the public with opportunities to comment on the draft policies orally and/or in writing, and incorporate public comments into the final policies. For any technology already in use but lacking such policies, law enforcement departments should immediately implement clear public policies that detail the parameters, requirements, and conditions of use.

Tech companies and developers of data-driven policing technologies have asserted trade secret evidentiary privileges as reason to deny defense discovery requests and subpoenas.[58] To facilitate transparency and avoid the exclusion of highly probative evidence,[59] companies that create and supply data-driven policing technology must waive, or otherwise not assert, claims of "trade secret privilege"[60] and must disclose the methodologies used to build the technology to law enforcement, the impacted communities where law enforcement departments intend to deploy the technology, the legislative bodies that represent the impacted communities, and the attorneys within the jurisdiction who specialize in criminal defense and civil liberties to ensure that the technologies are scientifically sound, are employed as intended, and are limited in scope to meet the articulated law enforcement need.

Any data-driven policing technologies that are used should undergo validation studies that allow them to be subjected to a *Daubert* or *Frye* analysis. As matters of constitutional due process rights guaranteed by the Fifth

and Fourteenth amendments, all individuals must be notified of their presence on data-driven databases that law enforcement departments access and utilize, including gang databases, strategic subject lists, and other data collected through social media monitoring. These individuals must also be provided the opportunity, through a private attorney or, if they cannot afford an attorney, an appointed attorney, to challenge their inclusion on such databases, the data accumulated from the databases, and law enforcement's interpretation of the data, as well as to seek removal from the databases.

All individuals, in accordance with constitutional due process rights guaranteed by the Fifth and Fourteenth Amendments, must be notified of their removal from any data-driven databases that law enforcement departments access and utilize, including, but not limited to, gang databases and strategic subject lists.

4. Race Equity

The analysis that jurisdictions undertake when considering whether to adopt any data-driven policing technology must be conducted through a race equity lens and include a racial impact statement. The "racial equity impact assessment"[61] must be conducted by experts trained in institutional and structural racism, as well as the history of racialized policing. These experts should work with legislators, law enforcement, and community members to examine the racialized impact of the proposed data-driven policing technology. If the racial equity impact assessment of the proposed data-driven policing technology concludes that use of the technology would harm the impacted community, the technology should be prohibited from use by law enforcement.

5. Accountability

If, through the processes detailed in Recommendations #2, #3, and #4, data-driven policing technologies have been approved, law enforcement departments must adopt and issue written protocols ensuring integrity and accountability, to ensure that the departments and the impacted communities can continuously monitor and otherwise gauge the use and effectiveness of these technologies.

Integrity and accountability measures must include data-keeping, annual departmental reports on the use and accuracy of the technology, measuring and evaluating the effectiveness of the technologies through auditing, and, based on the results of these accountability measures, determining whether the use of the technology should be modified or discontinued. All reports, evaluations, data, and accountability measures produced in relation to data-driven policing technologies should be made available to the public.

6. Resources and Access for Defense Attorneys

In accordance with the constitutional rights to discovery and confrontation guaranteed by the Sixth Amendment, prosecutors must provide to defense counsel notice and a description of data-driven policing technology that law enforcement has employed or has otherwise relied upon in the case, as well as any data based on the technology that the officers relied upon, assessed, or otherwise used in relation to the accused, including *Brady* material and any other data accumulated against the accused. Defense counsel must then be afforded time and resources to engage experts to analyze and interpret the data.

Defense lawyers must receive notice and training regarding the data-driven policing technologies employed by law enforcement departments in their jurisdictions, including the federal and state constitutional rights implicated by the technologies.

Defense lawyers should collaborate with other attorneys, technologists, and experts who understand the data-driven policing technologies employed against their clients, and should seek to incorporate law enforcement's use of the relevant tool(s) against their clients into all aspects of their representation.

Defense lawyers must have access to data-driven technology experts who can break down the technologies and consult on defense strategies vis-à-vis the data-driven tools that law enforcement relied upon to suspect, surveil, approach, or arrest, or otherwise employed against the accused.

Resources for public defenders and court-appointed counsel must be increased to respond to data-driven policing technologies in order to meet their constitutional obligation to provide zealous representation to clients impacted by these technologies.

7. Courts

Courts and prosecutors must be trained annually on the data-driven policing technologies employed by law enforcement departments, including the federal and state constitutional rights implicated by the technologies.

Judges must assess the reliability of a data-driven policing technology employed against the accused before determining whether it justified a Fourth Amendment intrusion. Data-driven technology must not form part of an officer's calculation of reasonable suspicion, unless the technology can be shown through typical evidentiary burdens that it is reliable.

Law enforcement authorities cannot utilize or otherwise rely upon data-driving technologies, such as gang databases, in any way that infringes upon the right to association guaranteed by the First Amendment.

8. Children and Youth

State and local jurisdictions must enact laws, policies, and protocols that protect the federal constitutional rights, state constitutional rights, and dignity interests of children and youth who are implicated or otherwise at risk of being criminalized by data-driven policing technology.

Law enforcement authorities should not include children under the age of 18 on any law enforcement database, or otherwise accumulate or access data specific to children under the age of 18 through social media monitoring or other data gathering practices.

Young people between the ages of 18 and 25 are especially vulnerable, disproportionately included on data-driven policing databases,[62] and therefore must be provided notice of their presence on any databases that law enforcement departments access and utilize, including gang databases, strategic subject lists, and other databases that incorporate social media monitoring. Individuals must be provided the opportunity, through a private attorney or, if they cannot afford an attorney, an appointed attorney, to challenge their inclusion on such databases, the data accumulated, and law enforcement's interpretation of the data, and, also, to seek removal from the databases.

An individual's ability to challenge their designation and inclusion on such databases, the data accumulated, and law enforcement's interpretation of the data should be ongoing, particularly given the impact of law enforcement interactions with children and youth on their personal development, self-esteem, and educational outcomes — including school attendance, suspensions, expulsions, and matriculation — and the correlation between these factors and involvement with the juvenile and criminal legal systems.

Any data, records, or other information contained in any law enforcement database through any data-driven policing technology and/or social media monitoring should be sealed and purged when the individual reaches 25 years of age, at which point the adolescent brain is fully formed.[63]

INTRODUCTION

The United States finds itself at a watershed moment for policing. Sparked by the continued tragic killings of unarmed Black men, women, and children by police officers – and particularly the galvanizing protests, demands, and efforts that have carried forward from the killings of George Floyd and Breonna Taylor – calls to reform, transform, defund, dismantle, and abolish law enforcement have taken center stage throughout the United States. The families of victims who have been killed by police officers, as along with other impacted residents, activists, civil rights organizations, elected officials, government agencies, and policing professionals, among many others, are pressing the need for radical change. Journalists and news organizations are reporting on and amplifying the voices for transformation.

These actions and efforts have crystalized in response to policing practices that have violated and hyper-criminalized individuals and communities of color. Over the last several years, U.S. Department of Justice investigations have revealed what residents of Black and Brown communities have long known and expressed: that abusive and unconstitutional policing pervade their lives. This is signified by the police violence that has taken, or otherwise jeopardized, the lives of men, women, and children of color; the lack of accountability for officers who have killed, injured and otherwise abused; legal doctrine and rules that protect such officers; and a criminal legal system that is omnipresent in the lives of individuals, families, and communities of color. Because of these conditions, which have evolved over the course of 402 years, the need to transform law enforcement, as well as the culture of criminalization, has long been urgent.

This Report discusses and analyzes one aspect of policing that calls for thorough examination: data-driven policing. In recent years, police departments have been turning to and relying on rapidly developing data-driven policing technologies to surveil communities, track individuals and, purportedly, predict crime. These technologies include algorithmic decision-making that departments claim can predict where crime is likely to occur, who will likely commit crime, and who will likely be a victim. These algorithms are thus designed to interrogate massive troves of data gathered in a myriad of ways, using inputs that can range from police-generated crime reports to publicly available social media posts. The outputs are then used to make critical decisions about patrols, or to make life-altering designations about individuals.

Included in this practice are law-enforcement databases, which police officers use to heighten levels of suspicion against individuals on the databases, to surveil those individuals, and, in various ways, interfere with their lives. Perhaps the most prominent of these databases, so-called "gang databases," are comprised of suspected gang members as well as individuals suspected by law enforcement suspect to be affiliated with gang members. A related component of data-driven policing are the various tools that police departments use to monitor and surveil individuals, groups, and communities who communicate with, or are otherwise connected to, social media.

The Task Force chose to use the term "data-driven policing" to refer to the various technologies and practices that were studied in order to avoid the shifting sands of terminology. In doing so, the Task Force understands that other technologies that do not utilize automated decision- making systems can fall under this broad

terminology, but for the purpose of this report data-driven policing is used to refer to the tools that analyze data to determine where, how, and who to police. In the time that the Task Force has started studying the issue, Palantir, a company that contracts with law enforcement agencies in major cities across the country, never described its tools as "predictive policing." The company instead advertises as a "data integration and analysis software platform."[64] As the term predictive policing fell out of favor, NYPD rolled out "precision policing," which "combined predictive policing analysis and community policing."[65]

The company that incorporated predictive policing into its name, PredPol, recently changed its name to Geolitica, explaining "[t]his new name – a mashup of 'geographical analytics' – better represents the direction our company has taken over the last few years."[66] What all of these tools and practices have in common is that they use algorithmic tools to process large amounts of data in order to focus law enforcement activity, be it where they patrol, who they patrol or who they find suspect. This report intends to address that practice, regardless of the terminology attached to it.

The Task Force also recognizes that expertise in the area of policing and expertise in the area of technology are often segregated. Experts tend to have a background in either policing or in technology, but not both. What became clear in the course of the Task Force's factfinding was that the technological tools being used to organize and interrogate police data are the natural extension of tactics that have been employed over a century of policing. The use of datasets collected and organized by law enforcement agencies has long-informed police actions, and those datasets reflected the implicit bias and explicit racism of the institution of policing. Using algorithmic tools to interrogate massive troves of police data does not correct for biased or racist policing practices; rather it entrenches it. As such, this report analyzes the technological tools that are guiding modern day policing practices in the historical context that birthed them.

Examining data-driven policing goes hand-in-hand with the efforts to "radically reimagine law enforcement"[67] in the United States. Some have raised the point, and the concern, that if efforts to slim, defund, and dismantle police departments are successful, law enforcement will rely even more on technology to, ostensibly, fill in the gaps. This line of thinking often fails to acknowledge the steep cost of these technologies, especially as compared to their unreliable outcomes. Even in the scenario where technologies are provided without cost by companies or purchased with third party funds, their impact will be at cross purposes with reform efforts.

Therefore, the need to understand and examine data-driven policing technologies is timely and necessary, particularly as law enforcement's use of such technologies continues to be largely unknown to the individuals whom they are surveilling or otherwise monitoring, the communities directly impacted by the deployment of these tools, individuals accused of crimes, their attorneys, and the courts.

To learn about, analyze, and respond to the myriad of issues that arise from data-driven policing technologies, the National Association of Criminal Defense Lawyers (NACDL) formed the NACDL Task Force on Predictive Policing in 2017. The Task Force's charge was two-fold. First, to examine the rise, implementation, and ramifications of data-driven technologies in policing. Second, to offer recommendations focused on whether police departments should utilize these technologies to serve their missions to reduce crime and enhance safety and, if so, how to do so in ways that ensure transparency, legitimacy, effectiveness, fairness, and due process to individuals charged with crimes or otherwise suspected or accused of criminal activity.[68]

From 2017 to 2019, the Task Force held meetings in Washington, D.C., Chicago, Los Angeles, San Francisco, and New York City. At these meetings, the Task Force met with a diverse array of witnesses, including technologists and industry experts, law enforcement personnel, academics, attorneys, advocates, and community stakeholders. The purpose of these meetings was to learn about the different types of technologies that police departments currently use or have used in communities throughout the country, and the impact of these technologies on individuals, families, communities, and the criminal legal system.

Witnesses offered their perspectives on the benefits and burdens of these technologies, and Task Force members posed questions to probe and flesh out the matters discussed. To preserve the information conveyed, the Task Force employed court reporters to transcribe the proceedings.[69] In addition to these in-person meetings, the Task Force met with a number of experts and stakeholders virtually and by telephone.

Notably, the Task Force attempted to speak with representatives from police departments in each city where the meetings were held – Washington, D.C., Los Angeles, New York, Chicago and San Francisco. Most affirmatively declined to meet, or were otherwise not available. To gain deeper understanding of policing technologies and to prepare for these meetings, the Task Force also read extensive literature on data-driven policing technologies, including but not limited to, books, academic articles, studies, reports, and news articles.

The Task Force's overarching recommendation is that police departments should not use data-driven technologies to predict who will commit crimes or may become the victim of crime, or where crimes will occur, and should not utilize databases, such as "gang databases," that lead to increased surveillance and criminalization. These "predictions" are ineffective, lack scientific validity, reinforce and exacerbate the racial biases that are inherent in traditional policing and all stages of the criminal legal system, and divert resources away from pressing community needs. Thus, the harms caused by these "predictions" far outweigh any claimed benefits. Also, "gang databases" are over-inclusive, hyper-racialized, and non-transparent. As with "predictive" technology, these databases exact much more harm on individuals, families, and communities than any claimed societal benefits.

The Task Force understands the reality that data-driven policing technologies, despite the racial costs and harms exacted, have already been implemented in jurisdictions across the country. Indeed, each stage of the criminal legal system – from decisions made at the bail and sentencing stages to predicting the "risk" of releasing individuals on parole – is increasingly reliant on technological tools that utilize algorithms to query the data and make recommendations based on "predictions" of future behavior or inherent criminality.

In turn, these "predictions" guide decision-makers, including judges and probation/parole/pretrial officials. Likewise, "big data" has become a featured component of modern policing, and will likely not only remain a mainstay but will become more pervasive. Thus, the Task Force understands that, despite our concerns, as well as the concerns and legal challenges raised by a wide swath of individuals and organizations, in some jurisdictions, the proverbial train has left the station.

As such, the Task Force sets forth additional recommendations specific to the use, implementation, and monitoring of data-driven policing tools. These recommendations focus on transparency and accountability, from the moment law enforcement intends to use these tools through when a prosecutor relies on the information gathered from to charge and prosecute an individual in court. Thus, the recommendations are focused on requiring that impacted communities be informed of law enforcement's desire or intent to adopt any data-driven tools that would impact those communities.

Most critically, this Report will serve as a guide for defense lawyers in jurisdictions where these tools are proposed or already in use. Defense attorneys must be aware of the existence and nature of these tools, the ways that police departments use them, the constitutional rights impacted, and how to account for these tools when representing clients. This Report offers some areas of focus for defense attorneys.

Residents of impacted communities and other stakeholders should be given opportunities to offer input on whether or not the technology should be adopted, and local governments must assess requests and not approve law enforcement's implementation of the technology without community input and agreement. In the event that such technology is employed, law enforcement's use of the technology must be independently audited to ensure that it is being used as represented, and that it is not undermining the legal rights, dignity, and humanity of the individuals and families within the respective community. The Task Force highlights these recommendations here and places them first in the recommendations section because the individuals and communities most directly and disproportionately impacted by all aspects of policing and each stage of the criminal legal system must work collaboratively with elected officials to determine what strategies best promote community security.

The Task Force sets forth several additional recommendations regarding the use and implementation of data-driven policing technologies. In addition to transparency, these recommendations are focused on the following themes: race equity; accountability; integrity; prosecutorial obligations to defense counsel; building knowledge, expertise, and capacity among the criminal defense bar, including public defenders and

court-appointed counsel; building knowledge, expertise, and capacity among courts; youth; and constitutional rights.

The four purposes of this report are:

1. to shed light on the rapid development and deployment of data-driven policing;

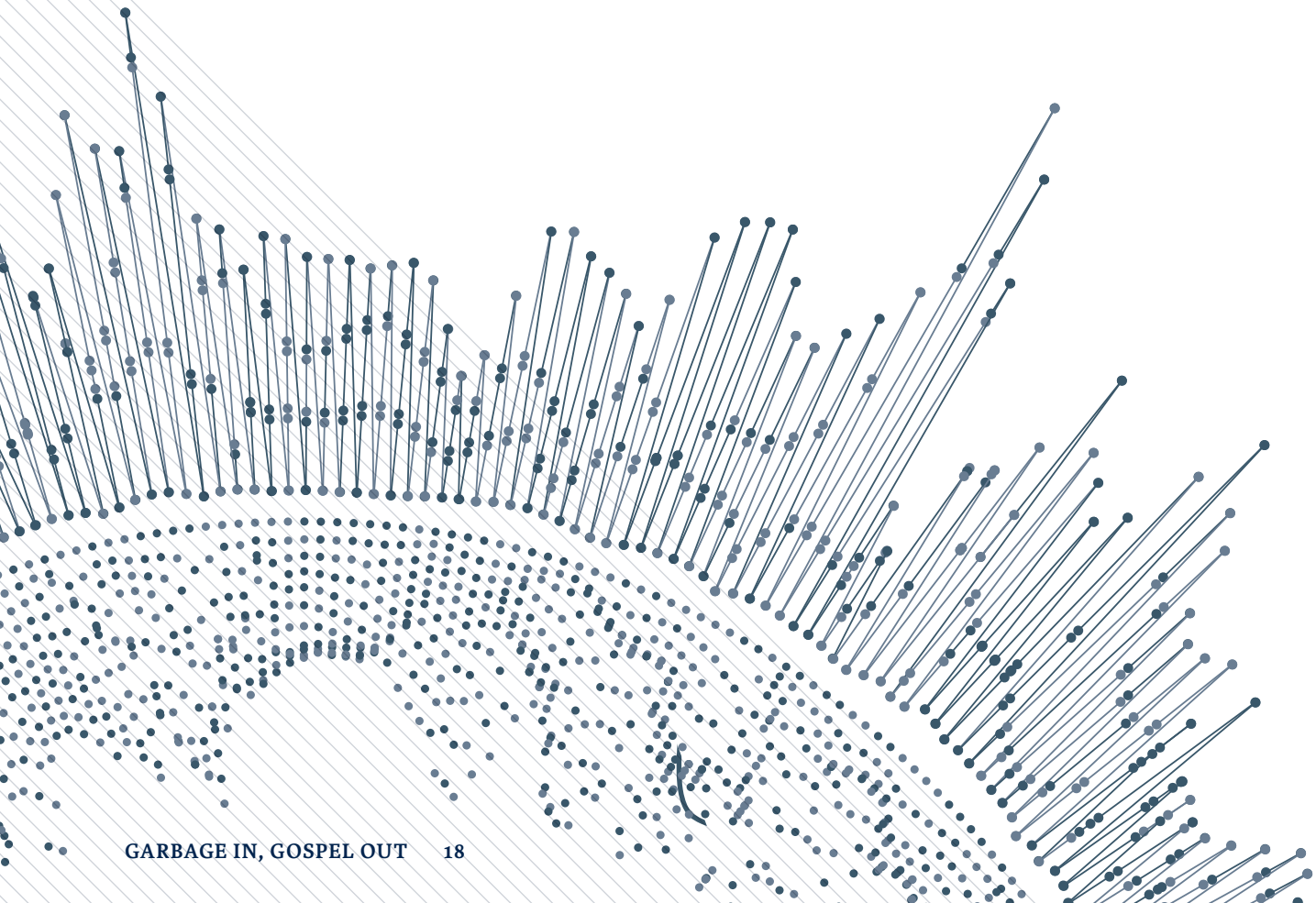
2. to explore the contours of data-driven policing in the context of the history of policing, the constitutional rights afforded to all people in the United States, and the criminal legal system;

3. to offer recommendations regarding data-driven policing to defense lawyers, lawmakers, community stakeholders, police departments, courts, and prosecutors that respond to the reality, enormity, and impact of data-driven policing; and

4. to suggest strategies for defense lawyers to represent their clients zealously, in places where data-driven policing technology is employed.

This Report begins with a brief overview of the history of policing and the economics of punishment in the United States. It then explores the rise of big data policing, contextualizing it within the long and problematic history of data collection and data analysis in policing. Following this, the Report lays out the landscape of data-driven policing and its varying iterations. Finally, the Report analyzes the myriad issues with data-driven policing practices and offers recommendations.

Throughout the history of the United States, policing has been steeped in racism and served as a tool of white supremacy. Indeed, during the drafting of this Report, the country has been reeling from tragedies involving the killings of unarmed Black men, women, and children by law enforcement officers. Millions have marched as activists, community leaders, academics, attorneys, and lawmakers have proposed ways to reform, transform, and dismantle policing. The history of policing is central to understanding the issues as they unfold today, and to grounding efforts to reform, transform, and dismantle. So too, it is central to our examination and analysis of data-driven policing technologies.



A BRIEF HISTORY OF POLICING AND THE ECONOMICS OF PUNISHMENT

As James Baldwin once said, "history is not the past; it is the present"[70]– this rings true and loudly in policing. This Report begins with a brief overview of the history of policing in the United States. It does so because no institution or system operates ahistorically. Thus, the history of policing provides necessary context when analyzing any aspect of present-day policing.

Policing has been stitched into the fabric of the United States since its founding over 400 years ago. As both a practice and a profession, policing has evolved over the centuries. In the North, the precursor to policing was informal "night watch" patrols, which formed during the Colonial era. Volunteers signed up for these patrols, which were focused primarily on looking out for community members who were engaged in prostitution and gambling.

In contrast to the North, which relied less on the labor of enslaved people, the forerunners of policing in the South were slave patrols.[71] As the name makes clear, communities formed these patrols to exert oppressive power and control over enslaved people. These patrols quickly evolved from informal groups of deputized white citizens tasked with monitoring and controlling enslaved people, to formal, professional, and paid groups of men who patrolled towns each night to surveil enslaved people and communities, and to enforce curfew and vagrancy laws.[72]

The Civil War and the abolition of slavery in 1863 brought the formal end to slave patrols. However, these patrols, as well as the continued subjugation of Black residents in the aftermath of slavery and the Civil War, were precursors to the modern criminal legal system in the country.[73] At the time that slavery was abolished, 90% of Black people in the U.S. lived in the South. Fearing, or simply outright opposing, the newly legalized freedom of formerly enslaved people, "[w]hite southerners initiated an extraordinary campaign of defiance and subversion against the new biracial social order imposed on the South and mandated by the Thirteenth Amendment of the U.S. Constitution, which abolished slavery."[74] Of course, the Thirteenth Amendment created a huge loophole to the abolition of slavery: "Except as a punishment for crime whereof the party shall have been duly convicted."

Southern states exploited this language by enacting Black Codes, which were laws that hyper-criminalized and controlled freed Black people.[75] Examples of such laws were "insulting gestures" and prohibiting Black people "from keeping firearms and from cohabitating with whites."[76] The codes also sought to maintain the exploitation of Black people and their labor through criminal penalties, imposing vagrancy and enticement laws designed to drive ex-slaves back to their home plantations."[77] The Black Codes were formally repealed in 1866, with the beginning of Reconstruction. However, the reprieve from these harsh and racist laws was short-lived, lasting only until 1877, when Reconstruction ended.

In the post–Reconstruction era, southern states again enacted, implemented, and enforced broad and seemingly unlimited criminal and civil laws, which hyper-criminalized, incarcerated, controlled, isolated, and oppressed Black men, Black women, and Black children. Among these laws were vagrancy statutes, which, inter alia, made it a crime to be unemployed. While these laws, as written, applied to Black people and whites, enforcement was

an entirely separate matter. Black people charged with vagrancy were hauled into court, convicted, and punished. In contrast, whites charged with this crime had the exclusive option of taking an oath of poverty in court. Once they did, judges waived any potential punishment.

Southern states imposed harsh penalties for crimes that lawmakers believed Black people were more likely to commit. Among these crimes were arson, rebellion, burglary, and assaulting a white woman, all punishable by death. States also enacted laws that banned Black people from possessing firearms, making or selling liquor, and selling farm produce without permission from the employer. These laws led to the imprisonment of unprecedented numbers of Black men, women, and children.

Southern states also designed criminal legal systems to criminalize, incarcerate, and otherwise punish Black people in ways that strongly resembled, and in significant ways, replicated slavery. The most notorious example was the Convict–Leasing System, which flourished in the Post–Civil War Era and served as an economic catalyst for southern states and employers.[78] Once Black people were convicted and sent to prison, states leased them to "employers." The "employers," in turn, paid states for the prisoners and assumed responsibility for their housing. Through convict-leasing (as well as sharecropping) "once-devastated towns…were again able to call themselves the cotton capitals of the world, and companies like United States Steel secured a steady supply of unfree [B]lack laborers who could be worked to death."[79]

One can draw a direct line from this history to the hyper-criminalization of Black men, women, and children today, along with the economic incentives embedded in the modern criminal legal system. Police officers continue to serve as arbiters of who is introduced to the system and who is not. As such, law enforcement strategies and tactics that overly criminalize and capture poor Black, Indigenous and People of Color ("BIPOC") men, women, and children, continue. The outcome is a criminal legal system that disproportionately criminalizes, incarcerates, and otherwise punishes these same populations for financial gain.

A. THE EVOLUTION OF MODERN POLICING

Policing has changed and evolved since it began as a profession in the mid-1800s. This evolution is marked by the priorities of policing over time – and at particular times – as well as by the roles that police departments assumed in communities.

During the mid-1880s to the 1920s, political machines dominated policing. Local party ward leaders often picked police captains and sergeants based on political favor or reward, and officers were largely involved in providing social services. During the height of the Prohibition era, President Herbert Hoover formed the National Commission on Law Observance and Enforcement to investigate a variety of issues concerning the criminal legal system, including law enforcement. The Commission found many police departments to be corrupt, poorly trained, and ineffective. As a result, policing became independent of party ward leaders.[80]

The 1930s to the 1970s observed a new series of reforms, as police departments increasingly emphasized "professionalism," "law enforcement," and "crime control." The 1960s was arguably the most critical decade of this era. In the 1960s, the U.S. Supreme Court expanded due process protections for individuals suspected of committing and charged with crimes by incorporating much of the Bill of Rights to the states through the Fourteenth Amendment. These advances included the exclusionary rule for illegally seized evidence,[81] the right to counsel for those unable to afford an attorney,[82] and *Miranda* warnings.[83] All of these changes had a direct relationship to the way in which policing was conducted.

The 1960s were also a time of immense tumult. The Civil Rights Movement is often remembered for marches, protests, violence against marchers, murders of civil rights leaders and activists, and historic legal victories. It is also marked by significant moments in policing. One such moment took place on "Bloody Sunday," March 7, 1965, in Selma, Alabama. On this day, men, women, and children planned to walk from Selma to Montgomery to protest the death of Jimmie Lee Jackson, killed one month earlier by an Alabama state trooper following a peaceful protest march. As the marchers crossed the Edmund Pettis Bridge on their way to Montgomery, they were met by state troopers "wearing white helmets and slapping billy clubs in their hands."[84] The troopers

used clubs, bullwhips, and tear gas to injure, maim, and traumatize. The television network ABC interrupted its regular programming – it was airing the 1961 film, *Judgement at Nuremberg* – to report live on the police violence, displaying the horrific brutality that has since been seared into the American consciousness.

The 1960s additionally bore witness to the dramatic progression of the Federal Bureau of Intelligence (FBI)'s Counter Intelligence Program (COINTELPRO), which disrupted groups such as the Communist Party, the Ku Klux Klan, the Social Workers Party, the Black Panther Party, as well as various civil rights leaders. As Natsu Taylor Saito writes, "[w]hile constituting a particularly intense period of governmental repression of political dissent, the COINTELPRO era represents not an aberration but the logical outgrowth of the previous use of law enforcement agencies to suppress movements for social change, a process that is still at work in the laws and policies being enacted in the name of countering terrorism."[85] COINTELPRO was the largest domestic surveillance initiative of its time, designed to disrupt and suppress civil and human rights movements that were building power in the United States. Its underlying rationale continues in today's implementation of high-tech surveillance tools and programs.

With fights for equality, the signing of landmark civil rights bills, continued racial segregation and entrenched poverty, opposition to the Vietnam War, government-sanctioned violence rooted in race, and Black communities in flames, the 1960s were remarkably turbulent. By the end of the decade, elected officials and law enforcement officials perceived newfound power in BIPOC communities. They also perceived increased lawlessness.

The criminal legal system shifted in the 1970s to models rooted in incapacitation and deterrence. The turn to large-scale forms of punitive punishment was largely in reaction to the movements from the decade prior, and with the belief that, in light of national and local upheaval, rehabilitation was not effective. Individuals and communities without social and political capital had limited access to law enforcement services as well as little influence with policing strategies. Police departments were overwhelming white, were largely unresponsive, and often showed outright hostility to the needs of Black residents. As a result, the communities that needed compassionate, collaborative, and cooperative engagement the most received it the least.

In the 1980s and 1990s, police departments, along with the criminal legal system, focused on "zero tolerance,"[86] "broken windows,"[87] and an escalation[88] of President Richard Nixon's so-called "War on Drugs."[89] These decades featured the prosecution of Black men, women, and children for an endless array of violations, misdemeanor offenses, and felony offenses as well as, again, a dramatic and unprecedented rise in the numbers of individuals sentenced to prison. They also saw the onset of a host of civil legal penalties – known as "collateral consequences" – that attach to individuals based on their criminal convictions. In addition, police departments, through agreements with public schools and with funding by the federal government, placed officers in public schools, leading to the criminalization of children, disproportionately Black boys and girls, through "zero tolerance" policies that converted routine school disciplinary matters to arrests and criminal cases.

In the wake of the attacks on September 11, 2001, state and local police departments were treated as the frontline for preventing terrorist attacks, and increased collaboration with federal law enforcement agencies and the militarization of state and local police departments followed. Fusion centers[90] were established, integrating state and local law enforcement agencies with federal law enforcement and intelligence agencies, and military equipment and surveillance tools soon flowed into state and local law enforcement agencies through initiatives like the 1033 program.[91] Law enforcement began using massive stores of data to classify and criminalize people. This fusing of law enforcement agencies and missions along with the militarization of state and local law enforcement agencies were sold as anti-terrorism tactics but instead exacerbated the surveillance, criminalization and incarceration of BIPOC communities.

B. POLICING TODAY

Throughout the course of history, law enforcement's impact on racially marginalized communities – particularly Black communities – has been unrelenting, forceful and destructive. Law enforcement has a dominant role in the everyday lives of these communities: on the streets, on the roadways, in homes, and in public schools. Two

aspects of policing have remained constant over time. First, policing has been rooted in the control and criminalization of Black men, women, and children.

As Professor Paul Butler pointedly stated in testimony to Congress, "[T]here has never, not for one minute in American history, been peace between Black people and the police."[92] In this regard, law enforcement, as an entity and a profession, is inextricably intertwined with the criminal legal system and its outsized role in BIPOC communities. As a result, the distrust that marginalized communities have in law enforcement stems from its disregard for the wellbeing of these communities and the humanity of their residents, as well as its direct connection to a criminal legal system that has criminalized generations and, as a result, defined and confined lives.

Second, lawmakers, prosecutors, and courts have historically deferred to law enforcement strategies, tactics, and practices. Law enforcement approaches and actions have been presumed legitimate and lawful, except in those rare, exceptional circumstances where courts have declared them to be constitutionally infirm.[93]

Currently, there is easy access to immeasurably large troves of data feeding the various facets and models of data-driven policing. Police departments throughout the U.S. are relying on ever-expanding technology to query and classify massive data sets to determine where to prioritize policing resources, who to surveil, and how to surveil. Continuing the trend of previous eras, law enforcement is deploying technology in ways that continue to hyper-criminalize Black and Latinx adults and children.

The Task Force is also mindful that low income and BIPOC communities suffer disproportionately from violent crime. Many residents of these communities have called for community-based solutions, alternatives to the criminal legal system (such as restorative justice), dismantling law enforcement departments (through defunding mechanisms that would shrink police departments and expand non-law enforcement services, as one example), and prison abolition. Regardless of perspective, residents want and need real solutions to the multifaceted and overlapping issues that their communities experience. Among these issues are institutional neglect, residential segregation, unequal and segregated educational opportunities, inadequate and often non-existent health care, lack of public transportation, sub-standard and unsafe housing, food insecurity, unemployment and underemployment, lack of economic mobility, redlining, and trauma.

The history and circumstances described above bring us to the present. Because data-driven policing is the latest expression of the inequitable and biased history of policing, it must be analyzed in light of that context. Police departments in the digital age rely on data to devise strategies and tactics that further those disparities, including where to deploy officers, whom to surveil, and whom to encounter. These tools are marketed as neutral "race blind" technologies that negate racism and bias, but in reality, they do little more than serve to tech-wash the racist and biased data they are built on.

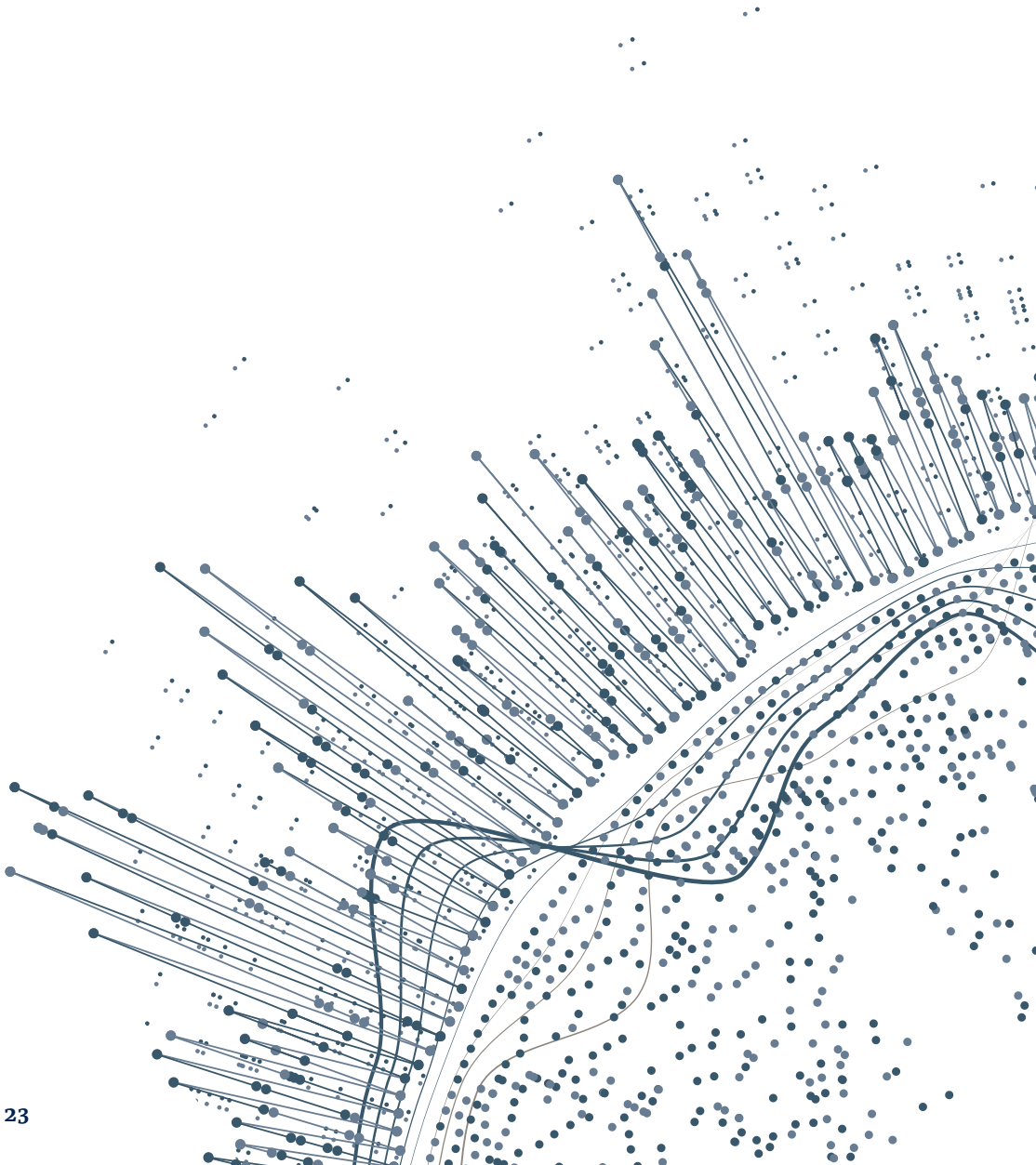
Such technologies cannot be divorced from the explicit and implicit biases that police officers bring to their patrols. Indeed, police feed the data that their officers have been collecting for decades into these tech tools. As such, data-driven policing tools have been shown to systemically replicate and amplify existing biases, both with regard to the information produced and an officer's actions based on that information. In addition, the companies that develop the tools control the process by which the data is analyzed, often with little transparency or insight into the process that delivers the results. In its current expression, law enforcement is essentially outsourcing critical aspects of policing to private entities that develop, market, and sell various types of data-analyzing technologies with little to no consideration of how the reproduction of existing police biases exacerbates the racialized nature of policing.

Private companies are developing the technologies that guide data-driven policing at a frantic pace. Today's technology will be outdated and replaced tomorrow. If there is one overarching, defining feature of data-driven policing technology, it is that much of it is deployed without notice or transparency to the public, and in particular, to impacted communities. Not only do communities often know nothing about the technology that police departments utilize against their residents, they often do not know that it even exists. Thus, law enforcement's use of the technology is subject to little or no oversight. The net result is that these technologies

implicate and undermine constitutional rights, as police officers use the data to suspect, surveil, accuse, search, and arrest, among other things, and prosecutors and other actors within the criminal legal system use the data to charge, prosecute, and adjudicate.

Overall, there are more questions than answers with regard to data-driven policing. The Task Force found a huge information gap between the developers who build the technology, the police departments that utilize the technology, the communities where the technology is deployed, the individuals who are charged with crimes connected to the technology, the defense attorneys who represent those individuals, civil liberties attorneys, and the courts.

As detailed in the Report, the Task Force concludes that data-driven policing is ineffective, costly, lacks scientific validity, replicates and exacerbates racial biases, hyper-criminalizes BIPOC individuals, families, and communities, and further disintegrates the already corroded relationships between law enforcement and communities. Thus, the Task Force's overarching recommendation is that police departments stop employing data-driven policing technology. Moreover, while current efforts to radically reimagine, transform, defund, dismantle or in other ways reduce the footprint of policing gain momentum throughout the United States, police departments should not rely on such technologies to fill claimed holes or gaps. In the alternative, the Task Force sets forth several recommendations focused on transparency, accountability, conditions of use, and fairness for individuals who are arrested, charged, and prosecuted based, in some part, on law enforcement use of data-driven policing technology.



BRIEF HISTORY OF SURVEILLANCE AND THE RISE OF BIG DATA

Contemporary data-driven programs in policing emerge from repeated attempts by criminologists, legal scholars, and law enforcement agencies to quantify and measure the complex social processes behind crime and disorder. Though these programs may vary in the types of data and techniques they employ, all of them are inevitably shaped by prior policing patterns, historical crime reports, and other records compiled by the police themselves; describing these technologies as "data-driven" obscures the discretion, biases and human decision-making inherent in the production of such data.[94] This section of the Report investigates how policing has historically functioned as a data creation practice[95] and the ways in which the generation and analysis of such data can be subjected to manipulation, distortion, and bias.

Though the UCR program encouraged the consistent collection of crime data, and could be considered somewhat dependable in the statistical sense, scholars have since pointed out that this approach leads to a phenomenon known as the "dark figures of crime," in which crimes committed by some groups are systemically undercounted due to the policies and practices of individual police departments.

A. THE ORIGINS OF CRIME DATA

Though data has always played a central role in surveillance by law enforcement[96], the widespread use of data collection and analysis in criminal proceedings did not begin until well into the late nineteenth century.[97] Early adopters of criminal statistics collected data on indictments, convictions, and acquittals using court records, but this was not a standardized practice across the U.S., with the exception of a few states.

After nearly a century of disorganization and confusion among the federal government's lawyers[98], the Department of Justice was established in 1870 to address an overwhelmed legal system following the Civil War. Under Section 12 of the Act to Establish the Department of Justice (ch.150, 16 Stat. 162), the Attorney General was required to "make an annual report to Congress...[on] the statistics of crime under the laws of the United States."[99] This piqued the interest of police chiefs at that following year's convention for the International Association of Chiefs of Police (IACP), which adopted a resolution calling for the formal compilation of crime statistics for police use in 1871. Historians have referred to this decision as the "first advocacy of uniformity in police crime records,"[100] representing a turning point[101] in the production, collection, and operationalization of crime data.

The IACP convened its first Committee on Uniform Crime Records (UCR) in 1927; law enforcement agencies began submitting data under a standardized UCR program based on crimes known to the police, as reported by the public or witnessed by members of law enforcement.[102] The formation of this new national crime data infrastructure ignited debate among members of law enforcement, criminologists, and statisticians surrounding the accuracy and reliability of police-collected data. Some argued that crime statistics should be based on the

records of criminal courts rather than collected directly by police departments,[103] while others like August Vollmer, then a police chief in Berkeley, California, claimed that such data was "a false index to crime because they may be variously interpreted"[104] and thereby could be subject to criticism.

Though the UCR program encouraged the consistent collection of crime data, and could be considered somewhat dependable in the statistical sense, scholars have since pointed out that this approach leads to a phenomenon known as the "dark figures of crime," in which crimes committed by some groups are systemically undercounted due to the policies and practices of individual police departments.[105] This was perhaps best exemplified by sociologist Donald Black's 1966 case study of police-citizen interactions in Boston, Chicago, and Washington, D.C., where he found that the decision to give official status to a crime was more likely the "outcome of face-to-face interaction between the police and the complainant, rather than a programmed police response to a bureaucratic or legal formula."[106]

The difference between crime data and actual crime reflects a longstanding observation about the police. Policing is not the passive collection of information, nor the identification of every violation of the law. Every action – or refusal to act – on the part of a police officer, and every similar decision made by a police department, is also a decision about how and whether to generate data. Crime data doesn't simply make itself known…[it] is the end result of many processes and filters that capture some aspect of the crime that actually occurs.

After observing over five thousand transactions between police officers and citizens, Black concluded that the behavior of individual police officers not only affected precinct crime rates, but also resulted in vastly "differential investigation of crimes, and hence, differential probabilities of arrest and conviction of criminal offenders,"[107] demonstrating that the rates of known crimes did not necessarily reflect the volume of citizen complaints that were recorded and included as crime data. Black's study arguably encapsulated the immeasurable influence of police discretion on the reliability of "crime data" and other official accounts of crime, particularly when the collection of such data has been shown to vary widely by race, class, and ethnicity.[108]

This skepticism towards the collection and use of crime data continued through the rest of the twentieth century, with ongoing claims that crime data was "at best a partial representation of crime in the community,"[109] and that such data could vary heavily based on police behavior. In fact, the former Chief of the Bureau of Criminal Statistics at the California Department of Justice, Ronald Beattie, even claimed in a 1941 article that police statistics were likely manipulated based on local political conditions,[110] often with a tendency to "report those facts which show a good administrative record on the part of the department."[111] By the mid-20th century, it became clear that police departments were increasingly calling upon quantifiably "measurable" crime rates as a "social fact, an empirical phenomenon with its own existential integrity,"[112] despite continued debate over what could be considered "objective data" within the context of the criminal legal system. As Elizabeth Joh further elaborates:

> The difference between crime data and actual crime reflects a longstanding observation about the police. Policing is not the passive collection of information, nor the identification of every violation of the law. Every action – or refusal to act – on the part of a police officer, and every similar decision made by a police department, is also a decision about how and whether to generate data. Crime data doesn't simply make itself known…[it] is the end result of many processes and filters that capture some aspect of the crime that actually occurs.[113]

In his meeting with the Task Force, Jay Stanley, Senior Policy Analyst at the American Civil Liberties Union (ACLU), additionally noted that there is still immense discretion for what is reported to and by the police, stating that "police have numerous ways, and in fact do, manipulate crime statistics for their own bureaucratic reasons."[114] For example, a 2014 survey of nearly 2,000 retired employees from the New York Police Department (NYPD) revealed that field officers were often pressured to manipulate index crime reports due to an over-reliance on crime numbers by police performance management systems, as well as pressures from leadership.[115] Despite the

prevalent use of such reports to measure police performance, the survey demonstrated a recurring trend of data manipulation in law enforcement, in which lower-ranking officers were less likely to accurately report crime and obey legal rules[116] when instructed to "produce" lower crime rates in their respective areas of patrol.

B. DATABASE POLICING

By 1967, the FBI established the National Crime Information Center (NCIC) as part of its "continued effort to develop a nationwide criminal records system,"[117] and as law enforcement agencies continued to generate, compile, store, and rely upon crime data from and for criminal investigations, the NCIC began to incorporate emerging computer technologies in 1971. Until desktop computers became widely available, police departments and criminologists had largely relied upon primitive techniques like mug shot cards and intelligence files to identify clusters of criminal activity.

Recent technological advances have since transformed the function and organizational structure of police departments. Between 1990 and 2003, the use of computers by law enforcement personnel skyrocketed from 5% to 56%, and by the early 2000s, the "vast majority of the nation's police agencies [were] using computerized data systems to monitor the activities of their officers (arrests, citations, calls for service, etc.)."[118] Contemporary policing scholars Stephen Mastrofski and James Willis wrote in 2010 that the growing presence of information technology (IT) in policing had become so normalized that it had given rise to a new form of surveillance called "database policing," where officers were effectively using computers and database management systems to "patrol" massive data files, looking for hits on information that they possessed on suspects.[119]

Law enforcement agencies were now capable of exchanging unprecedented amounts of data with a growing network of private and public sector risk-management institutions,[120] raising significant concerns about how database systems could potentially facilitate "pervasive institutional profiling" and differential treatment.[121] Further compounding matters, "the accuracy of databases is accepted as an article of faith, with courts according them a presumption of reliability,"[122] according to legal scholars Wayne Logan and Andrew Ferguson.

This concern was notably echoed by Justice Ruth Bader Ginsburg in her dissent in *Herring v. United States*, 555 U.S. 135 (2009), where she remarked that while "electronic databases [formed] the nervous system of contemporary criminal justice operations,"[123] insufficient monitoring and outdated information presented risk for error. It can be argued that the government was prioritizing the expansion of databases, rather than controlling for quality, with the understanding that data errors would inevitably prevail.[124]

Advanced methods in data-mining technologies and statistical modeling emerged in response to a growing need to bring meaning to vast amounts of raw data.[125] Prior to the development of contemporary data-mining techniques, traditional database analysis depended on the labor of individual database analysts, who would have to manually formulate queries[126] based on each particular database structure. This human-driven process was "slow, expensive, and highly subjective"[127] and could not keep up with the growing demands and complexities of data collection in both the public and private sector. The rapid expansion of databases in size and dimensionality[128] had created an unparalleled need for computer-automated data analysis methods.

At its core, data mining is the application of an algorithm that can identify and extract patterns in a database. This is accomplished through the selection and development of algorithms using "detailed domain-based knowledge and data familiarity in order to avoid irrelevant, misleading, or trivial attribute correlations."[129] These algorithms are often built to classify data into pre-existing categories.[130] Unlike traditional data-processing methods, data mining can transform "low-level data," which is usually too voluminous to understand, into "higher forms" of information and knowledge. These "higher forms" can be more compact, abstract, and serve a specific use, like a predictive model.[131]

Within the context of this Report, an algorithm[132] can broadly be defined as a "specified sequence of logical operations that provides step-by-step instructions for computers to act on data and thus automate decisions."[133] As explained by Daniel Kahn Gillmor, Senior Staff Technologist at the ACLU, in his meeting with the Task Force, algorithmic processes require that one takes "the complex field of information that is out there and reduces it to things that can be used as inputs and outputs."[134] As such, the techniques that are often employed in

data-driven policing typically fall under two types of machine learning paradigms: "supervised learning" and "unsupervised learning." Under "supervised learning," an algorithm can be taught to make predictions about future patterns using an initial training set that is designated by the developer.

As a result, these algorithms can introduce the assumptions and biases of their developers, though they have historically been used to identify individuals who are most likely to re-offend upon release, or to predict whether a given area will be "high-crime" or "low-crime" within a certain time period. Unlike supervised learning, where the developer's objective is to directly teach an algorithm using a provided data set, under "unsupervised learning," the algorithm is expected to teach itself to discover patterns without any reference data. Though this method may involve less of a direct influence from the developer, unsupervised learning can nevertheless still fail to consider the social context in which it is being deployed, as Andrew Selbst, et al. write:

> Machine learning systems are designed and built to achieve specific goals and performance metrics.... While performance metrics are properties of systems in total, technical systems are subsystems. Fairness and justice are properties of social and legal systems like employment and criminal justice, not properties of the technical tools within. To treat fairness and justice as terms that have meaningful applications to technology separate from a social context is therefore to make a category error.[135]

Though these emerging technologies have essentially enabled law enforcement agencies to pinpoint persons and activities on an unprecedented scale, some legal scholars have long argued that these data analysis techniques are, in fact, simply the computational automation of traditional investigative strategies in policing:[136]

> The application of data mining technologies to domestic security is the attempt to automate certain analytic tasks to allow for better and more timely analysis of existing datasets by identifying and cataloging various threads and pieces of information that may already exist but remain unnoticed using traditional means of investigation. Data mining can provide answers to questions that have not been asked, or even elicit questions for problems that have not yet been identified.[137]

The technological constraints of data storage and collection had historically prevented police departments from accessing criminal records beyond state borders. Modern-day policing is fueled by an almost unfettered access to immeasurable amounts of personal data,[138] ensuring that officers have seamless access to databases at the local, state, and federal level. According to a 2016 report published by the technology and justice non-profit Upturn, law enforcement agencies depend heavily upon third party vendors for brokerage and profiling products that use vast collections of both public and private data. This data can be used to build powerful custom-built search and mapping tools for police use.[139]

With no known studies on the accuracy of these tools and the data that support them, research suggests that these databases are often riddled with errors, and that "biases in the databases themselves, based on how data are collected, may also lead to disparate outcomes."[140] In fact, today's criminal intelligence databases are not only richly populated with "information about people who should be monitored and subject to scrutiny," [141] but are also frequently exempt from complying with the same constitutional and legal standards that govern criminal investigations. Even when information would normally require a warrant, U.S. law enforcement agencies at all levels can purchase access to commercial data brokers, without the need for a subpoena or a warrant, therefore doing an end run around Fourth Amendment protections. [142]

With such an unfathomably massive amount of data being constantly generated, early critics expressed concern over the possibility that police departments would use erroneous, irrelevant, or manipulated data. Like those skeptical about the collection of crime data in the twentieth century, critics have also argued that data mining had the potential to misrepresent "factual information in a false light."[143] Like researchers Solon Barocas and Andrew Selbst suggest, data mining, by nature, has therefore always been a form of "statistical discrimination"[144] because it holds the potential to place members of legally protected classes at systemic disadvantage under the guise of rational data science.[145]

This phenomenon of "statistical discrimination" is, in fact, particularly true for machine learning in criminal justice, as statistician William Isaac writes:

> These models are heavily reliant on the training dataset to estimate predictions. Machine learning algorithms are unaware, and in many cases, unable to adjust for institutional biases embedded within policing data. As a result, the presence of bias in the initial dataset leads to predictions that are subject to the same biases that already exist within the dataset. Further, these biased forecasts can often become amplified if practitioners begin to concentrate resources on an increasingly smaller subset of these forecasted targets.[146]

Although data-mining technologies, particularly within the context of policing and the criminal legal system, cannot singlehandedly determine human fates, the way they are designed and developed cannot be "value-neutral" as all technological systems reflect the values and interests of the humans behind the technology.[147]

C. CRIME MAPPING

Crime mapping, and more recently, geographic information systems (GIS)[148], come from a long history of attempts by criminologists to decipher the relationship between criminal activity and geographic locations, or "places." Crime maps even predate the rise of computers and the development of modern police administrations,[149] with some of the first recorded scholarship on crime mapping taking place in the 1930s under urban sociologists Robert Park, Clifford Shaw, and Henry McKay at the University of Chicago. The maps they developed had such an immense impact on the field of criminology that it gave rise to new theories on crime and place, which led to a reinvigorated interest in studying how technologies could help researchers understand patterns of criminal activity and served as a crucial precursor to contemporary place-based policing programs.[150]

If a group or geographic area is disproportionately targeted for unjustified police contacts and actions, this group or area will be overrepresented in the data, in ways that often suggest greater criminality.

Criminologists Paul and Patricia Brantingham[151] made similar advances in the field in 1981, positing that crime was "a complex event in which four things intersect at one time: a law, an offender, a target, and a place."[152] By shifting the focus in policing from people to the "multidisciplinary exploration of criminal events," such as crime sequences, clusters of crimes, and other environmental factors, the Brantinghams concluded that "places," not people, were the key element in crime, and that "it should be possible to predict the spatial distribution of crime and explain some of the variation in volume of crime between urban areas and between cities."[153]

Upon developing the ability to model the spatial and temporal distribution of crime, criminologists could identify crime concentrations, or "hot spots."[154] Scholars have historically defined hot spots as "geographically bounded spaces of varying size that are associated with heightened victimization risk and a proportionally greater number of criminal incidents than other similarly sized areas of the city." [155] Typically smaller in geography than entire neighborhoods, hot spots can comprise street blocks or street segments that experience "higher levels" of crime.

By the mid-1990s, "hot spot policing" had become a major priority for police departments and general crime prevention efforts, with governmental agencies like the National Institute of Justice (NIJ) specifically establishing a crime mapping unit in 1996 to "develop and promote criminal justice analytical tools using GIS technology."[156] Additionally, technological advancements in records management and dispatch systems to handle 911 calls for service allowed police departments to "systemically quantify" [157] the criminal activity that was occurring throughout a city.

Today, crime-mapping technologies have dramatically simplified the collection and analysis of crime statistics, and can be broken down by location, type of crime, and time period, among other factors.[158] With the ability to compare both historic and current patterns of crime, contemporary mapping tools allow police departments to identify hotspots based on recent crime patterns, redraw policing boundaries, and to even "connect with other jurisdictions to see how crime from one area affects neighboring areas."[159] In response to such advances in crime-mapping technologies, researchers have discovered that the underlying mathematical models are susceptible to "runaway feedback loops, where police are repeatedly sent back to the same neighborhoods

regardless of the actual crime rate" as a byproduct of biased police data. As Rashida Richardson writes:

> If a group or geographic area is disproportionately targeted for unjustified police contacts and actions, this group or area will be overrepresented in the data, in ways that often suggest greater criminality.[160]

By tech-washing racially biased policing practices and hiding behind data-driven tools that collect, use, and produce skewed data, law enforcement agencies are able to justify increased policing and surveillance in historically over-policed communities under the veneer of technological neutrality and objectivity; thus, crime mapping perpetuates a self-reinforcing cycle of bias and inequity.

D. SURVEILLANCE AND BIG DATA TODAY

1. Place-Based Crime Forecasting

Beginning in the 1990s, the most dominant model in criminology and policing literature has been tied to the development of advanced crime analysis techniques,[161] in addition to a new style of police management and organizational hierarchy that emphasized the use of crime data and criminal intelligence. Increasing interest in potential policing strategies that could "prevent" future crimes before they occurred paved the way[162] for the launch of data-driven programs like CompStat in 1994. Introduced by then-Commissioner William Bratton at the New York City Police Department (NYPD),[163] CompStat was one of the first performance measurement systems designed "to track crime statistics and have police respond to those statistics."[164]

Twice per week, precinct statistics and crime problems were projected onto overhead screens during CompStat Crime Control Strategy Meetings, where crime data that was once only available to police departments after three to six months was now available to precinct commanders on a weekly basis. Digital hot spot maps could be generated to show how crimes and police activities were geographically clustered. Officers soon attributed declining crime rates in New York City to CompStat, in addition to the adoption of broken windows policing and stop-and-frisk tactics.[165]

While the city purportedly experienced a historical decline in crime rates during the 1990s, former NYPD officers recently revealed that they were often driven to make "'highly unethical' alterations to crime reports"[166] to reduce crime. Sociologists such as David Greenberg have even claimed that there was no indication that CompStat had any substantive effect on violent or property crime rates in New York.[167] Following CompStat's example, one of the most popular algorithm decision systems (ADS) in policing is "predictive policing," which can be defined as "applications designed to identify likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions."[168] The term "predictive policing" itself came into fashion following a 2009 symposium organized by Bratton, after which the "NIJ distributed a series of grants that funded predictive policing research"[169] at universities and police departments across the country.

By 2008, Bratton had left the NYPD to begin assessing "the viability of a more predictive approach to policing."[170] From 2008 through 2011, he worked with anthropologist Jeff Brantingham, then-LAPD sergeant Sean Malinowski, former Justice Department senior executive Craig Uchida, and researchers at the University of California, Los Angeles (UCLA) to establish "PredPol,"[171] which is currently one of the largest vendors of predictive policing systems in the country.

PredPol has been described by the company as "a parsimonious race-neutral system that uses 'only three data points in making predictions: past type of crime, place of crime and time of crime."[172] Under the supervision of Brantingham and other criminologists, the LAPD began to experiment with a predictive algorithm that specifically targeted property crimes in Los Angeles and Santa Cruz. According to Andrew Ferguson, this experimental computer algorithm was attempting to predict areas of potential criminal activity:

> The predicted areas were precise – usually 500 by 500 square feet – and forecast a particular type of crime. Police officers on patrol received highlighted maps and visited those targeted areas as often as practicable within their regular patrols. It was believed that increased police presence at the identified areas would disrupt the continued pattern of property crimes.[173]

Following PredPol, technology companies like Palantir,[174] Amazon,[175] Microsoft,[176] Motorola,[177] Lexis-Nexis,[178] and ShotSpotter[179] have also entered the multimillion-dollar predictive analytics market with their own tools and programs. While the most basic place-based predictive models rely on data collected by the police themselves,[180] such as reported crimes and crimes discovered by the police, other programs have gone on to incorporate factors as variable as payday schedules, seasonal variation, liquor store locations, and potential escape routes.[181] Today, place-based predictive policing has evolved to target a much broader spectrum of crime, such as robberies, shootings, and gang-related violence, by analyzing geographic vulnerabilities, precursor crimes, and temporal patterns.[182]

The novelty of predictive algorithms in policing was not the use of quantitative data, but rather the application of artificial intelligence (AI) to collections of data that were once considered too large in quantity and too complex for police departments to analyze.[183] Just as prediction has always been a fundamental part of policing, the move towards predictive policing in the 2010s was more of a shift in tools, rather than strategy.[184] Modern predictive algorithms rose to prominence by supplementing, rather than wholly replacing, existing police techniques and strategies[185] through their ability to analyze and assess greater quantities of data "more quickly than any individual officer, crime analyst, or department ever could."[186]

2. Person-based Crime Forecasting

Like place-based predictive algorithms in policing, person-based algorithms are intended to predict who will be involved in a crime.[187] Police departments rely on person-based data-driven tools to predict who is most likely to commit crimes, in addition to who is likely to become the victim of a crime. They then act on those predictions to suspect, surveil, and encounter. This system is facilitated through many of the government-operated databases that often contain highly personal identification information, such as photos, addresses, and descriptions of scars and tattoos.[188]

These databases include "gang databases," which are often localized within cities or similar jurisdictions, and assemble and maintain a broad swath of identifying data on individuals known or suspected to be gang members, or associated with gang members. Police departments have maintained that these databases are needed to reduce gang-related criminal activity through tracking individuals, as well as sharing information with other law enforcement agencies.[189]

While a purported aim of these databases has been to "chronicle every known and suspected gang member in a community,"[190] persons of color comprise the vast majority of individuals listed on these databases. For example, a 2017 study conducted by the University of Illinois at Chicago found that over 72% and 20% of the individuals who were listed as "gang-affiliated" by the Chicago Police Department were Black and Latinx men, respectively.[191]

Such databases have also been found to be riddled with errors and shrouded in secrecy, effectively allowing police departments to maintain highly confidential identifying data, "without even a pretense of reasonable suspicion."[192] Individuals can be certified as gang members or associates, and entered into a gang database, simply based on their appearance or location, often without being notified of their inclusion in such a database or given the opportunity to challenge that designation.[193]

3. Social Media Monitoring

Another facet of person-based crime forecasting in policing is social media monitoring. According to a 2019 study conducted by the Pew Research Center, seven-in-ten Americans now use social media to "connect with one another, engage with news content, share information and entertain themselves,"[194] presenting significant opportunities for law enforcement agencies to mine social media posts and surveil individuals and groups.[195]

Indeed, a report published in 2016 by the International Association of Chiefs of Police (IACP) found that over 96% of police agencies use social media in some capacity, suggesting that the most common use of social media is for criminal investigations. This can result in a range of activities, spanning from a manual search on a social media platform to installing software on a targeted individual's computer and analyzing their social media

data.[196] As media scholar Daniel Trottier has theorized, the use of such social media monitoring strategies presents a new paradigm for profiling and policing, particularly when social media and its products are "recontextualized" and used for purposes outside of their "intended purview"[197]:

> Traditional surveillance, such as a guard tower or closed-circuit television, involves a comprehensive view from above, whereas surveillance via social media offers…more ways of seeing more people, due in large part to the saturation of social media and its associated practices in people's everyday lives around the world. This enables greater and more robust social surveillance.[198]

Investigations by law enforcement officials on social media can potentially violate a user's expectations of privacy, as officers operating undercover can connect with individuals and access information that would otherwise require a warrant to obtain. As an increasing number of agencies depend upon manual and automated social media content analysis tools to monitor and surveil, studies have repeatedly shown that these programs lack accuracy and often fail to grasp the nuances and complexities of social media's highly contextual nature.[199]

4. Implications

An estimated 2.5 quintillion bytes of data are now generated every day.[200] If a gigabyte of data storage cost hundreds of thousands of dollars in 1980,[201] that same amount of storage now costs just pennies, can be managed easily, and accessed anytime, anywhere.[202] With the ability to digitize, track, and store nearly every aspect of information, from Internet searches and social media posts to cell phone calls and retail purchases,[203] the term "big data" emerged to describe the use of increasingly large data sets in data science and predictive analytics.[204]

Traditional surveillance, such as a guard tower or closed-circuit television, involves a comprehensive view from above, whereas surveillance via social media offers…more ways of seeing more people, due in large part to the saturation of social media and its associated practices in people's everyday lives around the world. This enables greater and more robust social surveillance.

Sociologists like Sarah Brayne have even characterized "big data" more broadly as an "environment" that is vast, fast, disparate, and digital, shifting the focus from the "features of the data itself, to the social processes that give rise to big data collection and analysis."[205] In such a big data environment, individuals are constantly contributing "to a growing trove of data as they go about their daily lives,"[206] and such data plays an ever-expanding role in the surveillance of individuals and communities, and in determining who, where, and when police officers should monitor, encounter, search, and arrest.

Data-driven technologies have assumed center stage in policing. Police departments now collect, sort, and categorize unprecedented quantities of data on individuals, groups, and communities. These tools allow police to mine information and draw connections and conclusions – correctly or incorrectly – that previously would not have been possible. One key lesson of this Report is that the various types of data-driven policing do not operate in isolation. Police departments do not rely on one type of data-driven tool, but rather deploy many tools that interact with each other.

Furthermore, data-driven technology has placed a technological veneer over the racialized impact of policing. While policing tactics remain largely unchanged, the decisions that guide modern-day policing tactics are increasingly being outsourced to private companies with no perceived obligation to publicly disclose details of how their tools work. This Report describes and discusses the data-driven technologies that were known and made available to the Task Force through witnesses and information gathered from articles and reports. Thus, the Task Force's review of the various data-driven policing technologies is not exhaustive.

While the Task Force met with representatives from some data technology companies as well as policing professionals, representatives from some other companies and police departments declined our invitations to meet. Accordingly, there are additional data-driven policing technologies, as well as companies that produce these tools, that remain inaccessible or unknown to the Task Force and, as a result, are not discussed in the Report.

THE LANDSCAPE OF DATA-DRIVEN POLICING

Today, data-driven policing encompasses the many surveillance technologies, tools, and methods employed by police officers to visualize crime, target "at-risk" individuals and groups, map physical locations, track digital communications, and even collect data on individuals as well the communities they patrol. This can include any approach that incorporates a clear reliance on information technology, criminology theory, and predictive methods in policing.[207]

According to a report published by the RAND Corporation, predictive methods in policing can generally be divided into four broad categories: (1) Methods for predicting crimes, or approaches used to forecast places and times with an increased risk of crime; (2) Methods for predicting offenders, or approaches that identify individuals at risk of offending in the future; (3) Methods for predicting perpetrators' identities by creating profiles that accurately match likely offenders with specific past crimes; and (4) Methods for predicting victims of crimes by identifying groups, or in some cases, individuals who are most likely to become victims of crime.[208]

There are typically four key stages in place- and person-based policing. The first stage is data collection, which can range from historical crime data to "more complex environmental data, such as seasonality, neighborhood composition, or risk factors."[209] The second stage involves data analysis based on the type of crime each department wants to target and the resources available at their disposal. The third stage is police intervention, which involves "distributing crime forecasts to commanders who use them to make decisions about where to deploy officers in the field."[210]

During this third step in the predictive cycle, patrol officers will often focus their time and resources on surveilling the people and places models suggest are likely to be involved in future crime. This is followed by the fourth and final step, "target response," in which law enforcement intervenes to either serve as a deterrent, prevent the crime from occurring, or lead to the displacement of the crime to a different area.[211]

In order to implement these methods, predictive policing algorithms employ a variety of machine learning algorithms. Since the developers of data-driven policing technologies often assert trade secret evidentiary privileges to deny public access to the inner workings of their algorithms, the types of machine learning used in such programs are relatively unknown, and because many of these tools built on these algorithms are relatively new, or are continuing to change alongside advancements in technology, all are "are relatively untested, with only a handful of studies, reports, or empirical validation across jurisdictions."[212] At its core, predictive algorithms in policing programs are the "data-driven incarnation"[213] of what criminologists have been attempting to achieve for decades: to analyze past events, infer broader patterns, and to then use those insights to "prevent" future crime.

A. PLACED-BASED PREDICTIVE POLICING

The methodologies employed by today's place-based predictive policing systems depend on the basic premise that crime is not evenly dispersed geographically; that is, certain places are expected to experience higher rates of crime for a certain period of time, when compared to other places.

1. PredPol

PredPol was initially designed to serve as a data-analytics command structure that could direct police attention and resources to specifically targeted areas of criminal activity.[214] While most well-known for popularizing the concept of "predictive policing," PredPol in its early stages had "all of the same characteristics of past crime pattern identification strategies" [215] that had already been in use for years. Under the direction of UCLA anthropologist and PredPol co-founder Jeff Brantingham and George Mohler, an academic who later became PredPol's Chief Data Scientist, PredPol shifted its focus in the early 2010s to three types of crime: burglary, automobile theft, and theft from automobiles. Legal scholar Andrew Ferguson explains this shift as an effort to reduce popular theories in criminology to data points and precise predictions:

> The theory behind Predictive Policing 1.0 can be traced back to the work of criminologists who found that certain property-based crimes tended to have ripple effects in neighboring areas. Like contagious viruses, these crimes spurred additional crimes in the area, because either the same criminals came back to commit them, or certain environmental vulnerabilities existed to encourage crime.[216]

The academic theory behind PredPol's predictive algorithm is a statistical modeling method used in the field of seismology called the Epidemic-Type Aftershock Sequence (ETAS).[217] In 2015, Brantingham, Malinowski, and Mohler, among others, claimed that the ETAS models could predict 1.4 to 2.2 times as much crime compared to a crime analyst equipped with the same information and hotspot mapping practice. In their findings, police patrols using predictions based on the ETAS model led to an average reduction in crime volume of over 7%, whereas patrols based upon analyst predictions showed no such effect.[218]

Prior to re-branding as Geolitica in March 2021,[219] PredPol had historically published the algorithm they employed, claiming that such efforts were taken to demonstrate transparency, and to uphold the privacy and civil rights of its users.[220] However, as statistician David Spiegelhalter writes, such efforts for "transparency" by technology developers often fail to be accessible and usable, as "transparency does not necessarily provide explainability – if systems are very complex, even providing code will not be illuminating."[221] Instead, Spiegelhalter argues in favor of the "trustworthy algorithm," in which interested parties can assess the reliability of an algorithm's claims, and where "explanation is provided at multiple levels and in multiple formats."[222]

$$\frac{\partial A}{\partial t} = B + \frac{\eta D}{4}\nabla^2 A - \omega A + \theta\omega\delta$$

The patented algorithm used by PredPol.[223]

In its latest iteration, PredPol relied on three data points to predict where and when crime will take place: historical crime data by type; the locations of past crimes (by address) in those datasets; and the dates and times those past crimes occurred.[224] Based on this data, the algorithm produced maps comprised of small, square grids highlighting areas of potential crime hotspots. These boxes, each one 500 feet by 500 feet, mark the locations of where and when crime is "likely" to occur in a given area,[225] with this "likelihood" including probabilities as low as 0.01 percent.[226] After these maps are distributed to beat officers, police departments can use GPS trackers to monitor whether the officers patrol within the boxes, in addition to the length of time that each officer spends in a location.

PredPol has also claimed to only collect and employ data that indicates "what kind of crime occurred, where [it occurred] by address or by latitude and longitude, and when [it occurred] by date and time."[227] While speaking with the Task Force, MacDonald stated that PredPol does not collect any data on the offender or the victim, including arrest data and conviction data, nor "anything about the underlying socioeconomic or demographic composition of the neighborhood"[228] being targeted, recognizing that biases impact, and even dictate, how officers utilize discretion. By excluding arrest data, Sean Malinowski, former Deputy Chief of the LAPD who

pioneered the use of PredPol in Los Angeles, also claimed that excluding the arrest data would minimize the occurrence of any biases or self-fulfilling prophecies in PredPol's predictions.[229]

According to Brantingham in his interview with the Task Force in 2019, the crime data typically utilized in predictive analytics programs can often be divided into two distinct categories: "risk factors" and "event histories." Risk factors are any type of data that is believed to be relevant to understanding factors that are not by themselves criminal in nature, such as weather patterns, abandoned buildings, and the distribution of bars in an area. Event histories are verified past crimes, either associated with a place, time, or person.

By making these distinctions, Brantingham claimed that as a place-based predictive policing program, PredPol relies solely on event histories to decipher where, when, and what types of crimes are going to occur in the future.[230] This practice aligns with PredPol's base model, which allegedly focuses on the phenomenon of repeat and near-repeat victimization,[231] or "near-repeat methods." According to PredPol, these methods "operate on the assumption that some future crimes will occur very near to current crimes in time and place – that areas recently seeing higher levels of crime will see higher crime nearby in the immediate future."[232]

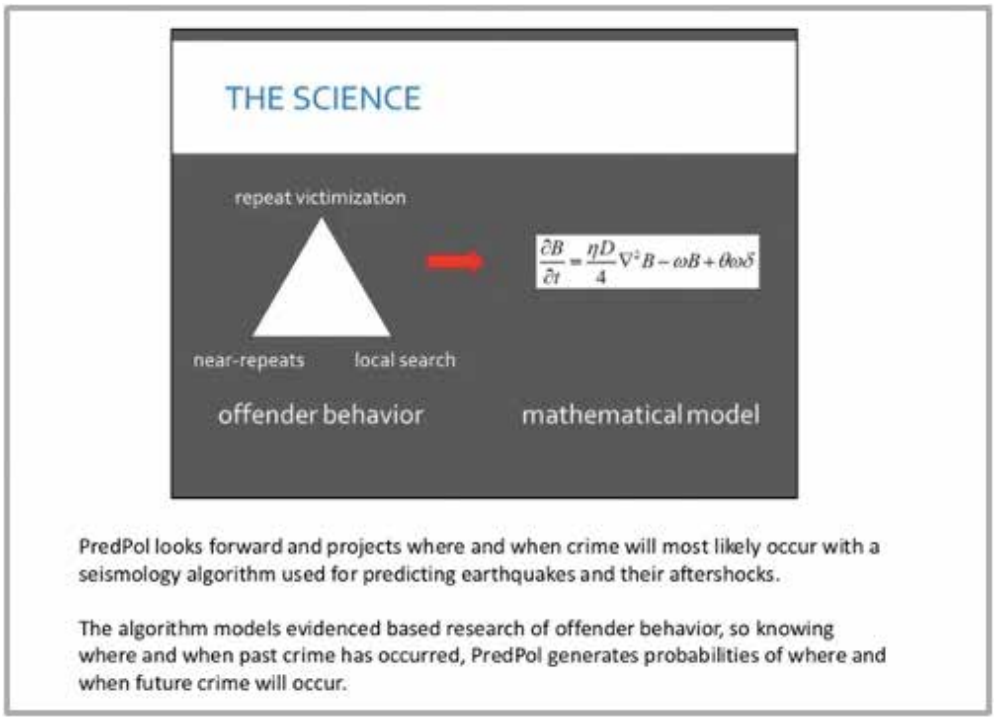
A promotional pamphlet from PredPol presented to the City of Columbia, SC in 2012.[233]

From the beginning, both the underlying theory and the initial experiments conducted by PredPol have focused on a limited number of property-based crimes as tied to place-based theories.[234] Unlike other crime forecasting tools,[235] predicting violent crimes or pinpointing individual criminals were not an aim of PredPol's early studies. In his meeting with the Task Force in 2019, Malinowski explained that as a place-based crime forecasting tool, PredPol's goal was always about "being in the right place at the right time."[236]

Santa Cruz and Los Angeles were the first two cities to pilot PredPol in 2012. Zach Friend, Second District Supervisor for Santa Cruz County and a former press information officer and crime analyst with the Santa Cruz Police Department (SCPD), supervised the testing in Santa Cruz. Friend provided PredPol with data on burglaries that occurred in Santa Cruz from 2004 to test how the algorithm predicted the location and frequency of burglaries in the following year. According to Friend, in his meeting with the Task Force, "within a quarter of a mile radius, [the PredPol algorithm] knew within a designed time frame of about an hour that a burglary would occur in a location, if predicted."[237]

For the SCPD, PredPol's value as a program was two-fold: not only did it let seasoned officers "know their intuition was right overwhelmingly," but it also gave newer officers the opportunity to "think about a situation

differently, critically."[238] Arriving at the height of the early 2010s recession, the program was an opportunity for agencies like the SCPD, which were "losing law enforcement officers functionally [while] not seeing a decrease in calls for service or a decrease in crime," [239] to supposedly compensate for a deficit in departmental resources.



A screenshot from a slide presentation titled "Predictive Policing Tacoma Overview Deck (2012 July)" *obtained via public records request from the Tacoma Police Department.*[240]

Though Santa Cruz was one of the first cities in the U.S. to adopt PredPol, it also became one of the first cities to enact a new ordinance banning predictive policing in 2020, with Santa Cruz Mayor Justin Cummins specifically stating that "if policing itself is biased, then the data that's informing those models will be biased."[241] This follows a trend of police departments cancelling their contracts with companies like PredPol due to budget constraints or dissatisfaction with the software.[242]

2. HunchLab

Designed by the Philadelphia-based GIS firm Azavea in 2005, HunchLab was described as a "proactive patrol management tool"[243] by Jeremy Heffner, former Product Manager and Senior Data Scientist at HunchLab, in his meeting with the Task Force. Most of its competitors were brought to the market by major corporations, such as IBM, Microsoft, Motorola, Hitachi, and LexisNexis, or were backed by venture capital and CIA seed funding. HunchLab quickly distinguished itself within the predictive analytics sector for not being tied to "shareholders, investors, or covert government funding schemes" when first founded.[244]

HunchLab was later sold to ShotSpotter, a gunshot detection technology company, in 2018, with Robert Cheetham, President and CEO of Azavea, citing issues with funding, project management, and resources.[245] HunchLab's platform has since been renamed "ShotSpotter Missions," and is now offered as an add-on product to ShotSpotter's gunfire detection solution, as well as a stand-alone application.[246] The Task Force did not study HunchLab's iteration under ShotSpotter.

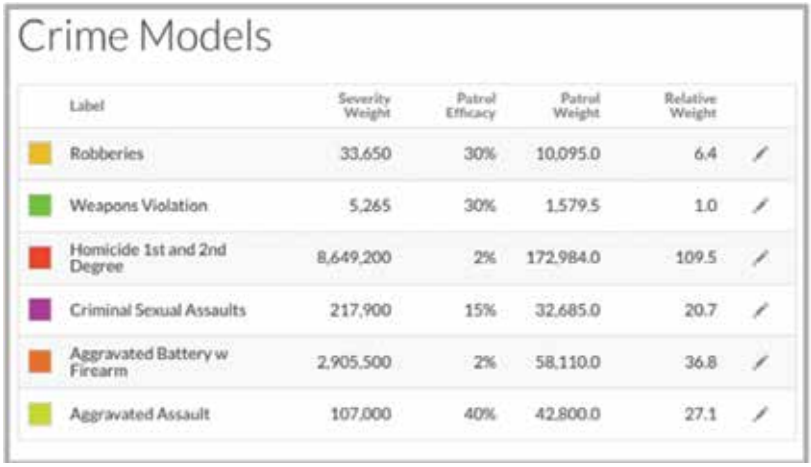
Unlike PredPol, HunchLab's program not only incorporated public reports of crime, but also the interaction of social, behavioral, and physical risk factors, such as weather patterns, moon phases, the location of bars and bus stations, and even the schedules of major sports events.[247] HunchLab's algorithm was trained on a set of training examples using the client department's crime data from the previous five years, in addition to incorporating

several non-crime-related data sets; all of which are then mapped onto a grid of 500 feet by 500 feet cells that cover a client's jurisdiction.[248] Contrary to other predictive policing programs, HunchLab's system generated separate predictions for each of the different models that a department configures for specific crime types. These individual predictions are then combined to create target areas based upon the crime weights set by each department. By assigning severity weights to each type of crime, departments could decide to prioritize certain crimes based upon the impact of the crime on the community in question.[249]

According to promotional materials provided to the NYPD and obtained through a public records request by the Brennan Center for Justice, HunchLab tested several approaches to the concept of "weighting" crime. The first was to use published information about the cost-of-crime, where weightings would be expressed as a monetary value; such numbers were typically only available for major crime types. Another approach was to use sentencing guidelines, though sentencing decisions are significantly racially skewed.[250] However, for HunchLab, these guidelines were utilized for the program's weighting system as "a measure of the import that society places on various offenses,"[251] as explained by Heffner in his meeting with the Task Force:

> [If] this configuration is set to forecast homicides, aggravated assaults, robberies, motor vehicle theft, theft of vehicles, and residential burglary, [then] the system needs to be told the value of preventing each of those incident types. This is used to weight them appropriately. This particular set of weights is from the RAND Corporation's cost of crime numbers, which tried to account for the societal community impacts of different crime types. I think a better weighting system would be that you take a community group, and you engage them in choosing the crime types that are going to be in the selection and the weights because a particular community might have different belief structures about what's important and what they want the police to address.[252]

Unlike published information on the cost-of-crime, sentencing guidelines were available for all types of crimes[253] and were not limited to a specified list of examples. Since HunchLab's focus for predictive policing was "to prevent crimes (and their associated arrests and incarcerations), by aligning proactive work with the areas where the most potential incarceration would occur due to crime,"[254] HunchLab claimed that their weighting system was essentially designed to "reduce the incarceration rate."[255]



An example of crime model weights as presented in HunchLab, in which command staff are able to set the crime priorities used by the system to generate predictive mission areas.[256]

B. PERSON-BASED PREDICTIVE POLICING

Person-based predictive policing attempts to identify individuals or groups who are likely to commit a crime — or to be victim of one — by analyzing for risk factors such as past arrests or victimization patterns.[257]

1. Chicago's Strategic Subjects List

At the forefront of person-based predictive policing was the Strategic Subject List (SSL): Chicago's contribution to the increasingly expansive and intrusive regime of predictive analytics,[258] in which an algorithm was used to identify individuals who were likely to be involved in future criminality. The SSL was developed in 2012 by Miles Wernick, then an engineer at the Illinois Institute of Technology (IIT), who had previously worked with the U.S. military and had been engaged in other predictive analytics work since the 1980s.[259] Several years prior, the National Institute of Justice (NIJ) had awarded the CPD with over two million dollars in grants to test and implement a new predictive analytics program. After entering a partnership with the CPD in 2009, Wernick and his team of researchers at IIT began working on a program that could allegedly identify networks in the city that were at risk of an "uptick in crime."[260]

According to CPD Special Order S09-11, the SSL algorithm ranks "individuals with a criminal record according to their probability of being involved in a shooting or murder, either as a victim or an offender."[261] Individuals are ranked with a score between 1 and 500, and the scores are re-calculated every day; they do not distinguish between a potential crime victim or a potential perpetrator. Closer attention is paid to individuals with scores of 250 or above,[262] though it is unclear why 250 was chosen as the threshold. The most recent and publicly available version of the SSL took eight factors into consideration: the number of times being the victim of a shooting incident, age during the latest arrest, number of times being the victim of aggravated battery or assault, number of prior arrests for violent offenses, gang affiliation, number of prior narcotic arrests, trend in recent criminal activity, and number of prior unlawful use of weapon arrests.[263] While these factors can appear race-neutral, all of the data results from areas that are already over-policed, resulting in a list comprised almost exclusively of Black and Brown people.

The CPD has also "not been forthcoming regarding the weight which the algorithm gives each of these eight factors when calculating a final score." Using linear regression analysis on SSL scores obtained by a freedom of information request, an investigation conducted by the *New York Times* was able to estimate the likely impact of each factor, concluding that age and victimhood were the two elements with the greatest impact on the SSL risk score.[264]

Factor	Marginal Increase in Risk Score
No. of assault or battery incidents (as victim)	+34
No. of shooting incidents (as victim)	+17
No. of arrests for violent offenses	+15
Trend in criminal activity	+14
No. of unlawful use of weapon arrests	+12
No. of narcotics arrests	+5
Gang affiliation	+4
Age (per decade)	–41

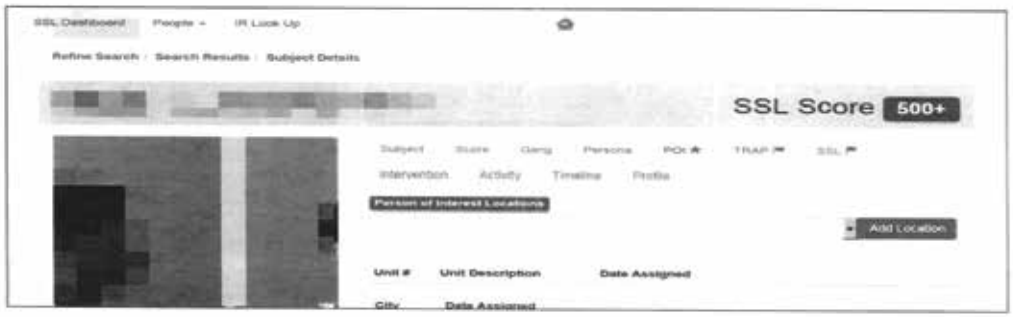
A chart demonstrating the likely impact of factors when calculating the final SSL score.[265]

As explained in his meeting with the Task Force, researcher David Robinson also found that the most important factor in an individual's SSL score was their age.[266] While the CPD have described the algorithm as one that does not use information about where the person lives, or race or gender, and uses "only the pattern of criminal activity,"[267] the database includes significant information on gender, race, residency, and recent arrests.

An early RAND study on the SSL program found that the SSL uses "social networks (in the form of co-arrests) to previous homicide victims to prevent the likelihood of someone becoming a victim of homicide."[268] Early pilots of the model focused heavily on literature examining correlations between victimization and the social connections to those who were victims of homicide.[269] Indeed, in a 2013 interview with *The Verge*, Wernick

drew comparisons between the spread of gun violence to that of a blood-born pathogen,[270] stating that, "people who know each other and hang out in the same circles – people who are a part of the same social network – infect each other with their interests, [and] when those interests include high-risk activities such as carrying or a gun or selling drugs, that leads to predictive trouble."[271]

The CPD has claimed that the SSL provided social services to those at particular risk of violence, including through the "Custom Notifications" program, more formally known as the Custom Notifications and Targeted Repeat-Offender Apprehension and Prosecution (TRAP) program.[272]According to CPD policy, the Custom Notifications program identifies at-risk individuals and reaches out to "advise them of the risks and consequences of their actions, should they engage in criminal conduct" with the goal of ensuring that the individual is informed of the "devastating impact of gun violence within their community."[273]



A screenshot of an individual's entry in the Strategic Subject List, obtained by the South Side Weekly in June 2017.[274]

Though the Custom Notifications program was supposed to present opportunities for intervention for those seeking assistance, researchers have found "little public evidence of those interventions actually occurring."[275] Indeed, the Custom Notifications process "fundamentally remains a law enforcement deterrence tool," with the notification including a description of potential enhanced federal and state sentencing options, as well as the potential for seized assets. In his meeting with the Task Force, activist Freddy Martinez noted that for individuals who were arrested after interacting with the Custom Notifications program, the prosecutor would "try and seek the harshest possible penalty."[276]

The SSL algorithm was inspired by the work of sociologist Andrew Papachristos, whose research focuses on how gun violence spreads within social networks.[277] Viewing recent advances in data analysis technology as a critical step towards improving public health and community relations, Papachristos envisioned a system in which social and medical providers could "provide immediate guidance to those in harm's way," and where service providers could reach out to young people "who would be better served through diversion as opposed to detention."[278] Since the CPD's implementation of the SSL, Papachristos has distanced himself from the program, citing concerns about its transparency and its fixation on identifying "offenders" in communities, which simply reinforces the ways in which "American devalues the lives of young people of color."[279]

Chaclyn Hunt, a civil rights attorney and the Director of the Youth/Police Project for the journalism non-profit Invisible Institute, found that CPD officers routinely used language associated with the SSL program to track and potentially target high school students:

> I work with high school kids, and a lot of them enter the criminal legal system at some point. I have kids who have been told by police officers, "'We know you're really likely to be shot"; [but] not that they're on the Strategic Subject List. They are using the language of this with a specific kid, after that kid has been arrested, [yet] none of my kids have ever reported being notified; nor do they have family or friends who have been notified.[280]

The SSL disproportionately targeted Black and Brown young men. After the CPD lost a lengthy legal dispute with the *Chicago Sun-Times* in 2017 and were forced to release a version of the SSL database using arrest records from August 2012 through July 2016, journalists found that 56% of Black men in the city ages 20 to 29 had an SSL score, despite claims that "resulting scores do not overestimate or underestimate risk for any specific

demographics."[281] In fact, of the nearly 400,000 people who are on the SSL, over 50% are Black and 25% are Latinx, 45% are under 30 years of age; and over 75% are male. Of the people labeled as "gang-affiliated" in the SSL, over 67% have never been arrested for a violent offense or the unlawful use of a weapon.[282]

The *Sun-Times* investigation additionally concluded that "nearly half of the people at the top of the list have never been arrested for illegal gun possession, and 20 of the 153 people deemed most at risk to be involved in violent crime, as victim or shooter, have never been arrested either for guns or violence."[283] Further research conducted by Upturn concluded that "more than a third of individuals on the list have never been arrested (133,474)," contradicting the CPD's claim that the list consists of only those with an arrest record.[284]

The SSL program was "dumped"[285] by the CPD in 2020 after a report published by the City of Chicago's Office of the Inspector General (OIG) concluded that the SSL had not been effective in reducing violence, and that "of the 398,684 individuals recorded in one version of the model, only 16.3 percent were confirmed to be members of gangs."[286] In her meeting with the Task Force, Jessica Saunders, formerly a researcher at the RAND Corporation, additionally noted that there was no evidence that any person-based predictive policing strategies like the SSL had proven "effective" by any metrics. In the case of the SSL, the CPD's use of the program had no effect on citywide violence levels, and the 2013 version of the program was, according to Saunders, not nearly as valuable as the department had previously claimed:

> We found that being on the list didn't change anything, except that it made it more likely that they were arrested for gun violence.[287]

The discontinuation of the program was immediately followed by the introduction of two new programs akin to the SSL, the Subject Assessment and Information Dashboard (SAID) and the Crime and Victimization Risk Model (CVRM), under the CPD's Special Order S09-11.[288] CVRM was already an active part of the SSL, but its newer version in SAID allegedly "relies on a shorter list of predictors, excluding previous narcotic arrests and gang activity."[289]

2. Palantir and the "Gotham" Program

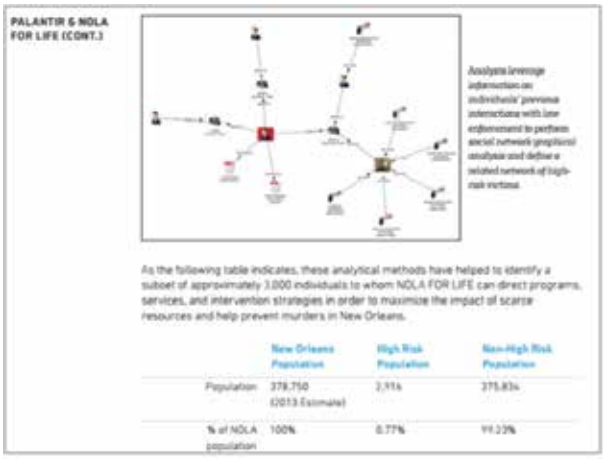
Palantir is one of the most secretive companies in big data analysis. Serving as "an information management service" for corporations and dozens of local, state, and federal agencies,[290] scholars have described it as a "secondary surveillance network" due to its extensive catalogs and networks of data across the country. Though Palantir's "Gotham" program began as a tool specifically developed for the needs of government institutions like the Department of Defense and the National Security Agency,[291] it has since been adopted by police departments to aggregate and synthesize data in such a way that "gives law enforcement nearly omniscient knowledge over any suspect they decide to surveil."[292] Oftentimes, Palantir's relationships with law enforcement were not known to the public, and at times unknown even to local lawmakers.

In New Orleans, defense attorney Kevin Vogeltanz was unaware that law enforcement used Palantir's Gotham program to designate people as gang members until a 2018 exposé ran in the The Verge.[293]After his client Kentrell "Black" Hickerson was convicted in 2016 of conspiracy, racketeering, and other gang-related charges, Vogeltanz accused prosecutors in the case of suppressing analytic evidence obtained through the use of Palantir, "arguing he had a right to view the evidence if Hickerson's name surfaced as being affiliated with a gang."[294]

Hickerson's case was the first in the Orleans Criminal District Court in which "the possible use of Palantir software has been pointed to by a defense attorney as potential evidence that should possibly be subjected to discovery rules."[295] Despite the NOPD's claims that Palantir's software played no role in Hickerson's specific indictment and prosecution,[296] Assistant District Attorney Alex Calenda allegedly "admitted during Hickerson's hearing that he was an 'end user' of NOPD's Palantir system."[297] According to an investigation conducted by *The Verge*:

> Palantir's prediction model and risk assessment database in New Orleans used an intelligence technique called social network analysis (SNA) to draw connections between people, places, cars, weapons, addresses, social media posts, and other indicia previously siloed criminal databases. After entering a

query term, such as a partial license plate, nickname, address, phone number, or social media handle or post, an NOPD analyst would review the information scraped by Palantir's software and determine which individuals are at the greatest risk of either committing violence or becoming a victim, based on their connection to known victims or assailants.[298]


PALANTIR & NOLA FOR LIFE (CONT.)

As the following table indicates, these analytical methods have helped to identify a subset of approximately 3,000 individuals to whom NOLA FOR LIFE can direct programs, services, and intervention strategies in order to maximize the impact of scarce resources and help prevent murders in New Orleans.

	New Orleans Population	High Risk Population	Non-High Risk Population
Population	379,750 (2013 Estimate)	2,916	375,834
% of NOLA population	100%	0.77%	99.23%

A slide from a Palantir presentation titled, "NOLA Murder Reduction: Technology to Power Data-Driven Public Health Strategies" obtained by The Lens.[299]

In his meeting with the Task Force in 2019, Vogeltanz added that Palantir's collaboration with the NOPD was a completely secret program, and that the Mayor's Office never informed the City Council of the program's existence.[300] At the beginning of Palantir's six-year-long contract with New Orleans, the company offered its services free of charge and gave law enforcement the ability to identify people deemed likely to "either commit gun violence or be the victim of it."[301] The data used by Gotham typically came from social media platforms, in addition to the NOPD's data on gangs, probation, parole information, calls for service, and "every documented encounter the NOPD has with citizens, even those that don't result in arrests."[302]

The program's data analysis would identify "names and connections between people on FIs [field interview cards], on traffic stops, on victims of reports, reporting victims of crimes together, whatever the case may be,"[303] and according to Vogeltanz, analysts would reportedly take advantage of such information to generate leads and to "predict, in a roundabout statistical way, who the most likely victims of violent crime were in the city."[304] This was confirmed in emails obtained by *The Times-Picayune* through a public records request in 2018, which revealed NOPD spreadsheets that ranked individuals based on the "number of gun related events (weighted according to severity) with which a person was associated."[305]

The database was additionally employed to "hold together criminal conspiracy cases," and to guide detectives toward potential leads in their investigations.[306] As Vogeltanz explained to the Task Force, an extensive amount of data from gang registries were often extracted to produce social networking graphs that would be used to "probe possible ties between people linked to gun violence in New Orleans"[307]:

> The entire sheriff's department arrest and booking registry was data mined and put into an Excel spreadsheet. It's your name, your race, the location of where you were arrested, the location of where the crime allegedly happened, any analysis that you have, any other people that were with you … a massive amount of data.[308]

New Orleans is not alone in having implemented data-driven policing technology without oversight or public notice. When public processes are circumvented, the public is left largely unaware and defense attorneys are denied key evidence in their clients' cases and the opportunity to challenge that evidence in court.

3. Los Angeles' Operation LASER

Funded with nearly a million dollars from the Federal Bureau of Justice Assistance, the Los Angeles Strategic Extraction and Restoration Program, a LAPD program commonly referred to as "Operation LASER," was first launched in 2011. By the end of 2017, twenty-one patrol divisions were using the program designed by Palantir Technologies to pinpoint likely criminal actors, and to develop "Chronic Offender Bulletins" that could identify targeted individuals.

First field-tested in the Newton Division of the LAPD, officers were tasked with identifying corridors called "LASER zones," where gun violence would likely occur and where the LAPD would subsequently increase patrols.[309] Operation LASER involved the use of both location- and offender-based strategies[310] linking computer systems at the LAPD interdepartmentally, and also with other local, state, and federal agencies. By accessing all of these previously independent data systems, the program was allegedly capable of directing officers to places where crime is most likely to occur, while also keeping track "of ex-convicts and others they believe are most likely to commit them through technology such as license plate scanners and cellphone trackers."[311]

Sociologist Sarah Brayne, who conducted months of field work at the LAPD and other law enforcement agencies in the area, noted in her meeting with the Task Force that Operation LASER was a pivotal example of police departments shifting towards more quantifiable intelligence gathering tools in the age of big data:

> Once somebody comes under suspicion, you can then retroactively draw on all of these data points. There's this idea of building up networks and dragnet surveillance where you collect information on everyone rather than merely individuals that are under surveillance by law enforcement suspicion... I call this the secondary surveillance network, where people do not need to have any direct law enforcement contact, but they are in the corpus, they're in this law enforcement data. If they do ever come under suspicion, all of that information on them can be leveraged.[312]

Brayne's research began in 2013, when the LAPD had signed contracts with three programs: Palantir's Gotham, PredPol, and Operation LASER. Brayne details an interview with a captain whose goal was simply to "get people 'in the system': to capture larger and larger amounts of data on seemingly harmless individuals in the hope that the data would help solve a crime later on."[313] According to Brayne, when officers came into contact with someone who seemed "suspicious," they were instructed to fill out a field interview card with the person's name, address, physical characteristics, vehicle information, gang affiliation, and criminal history. These field interview cards recorded "information not only about the individual in question, but also information on people the individual is with."[314]

Like the implementation of Palantir's Gotham program in New Orleans, the LAPD focused on identifying and "scoring" individuals based on past crime and arrest data. The "Chronic Offender Bulletin" generated by Gotham in Los Angeles was routinely provided to police for surveillance and investigation purposes with the basic premise that they were to "target with laser-like precision the violent repeat offenders and gang members who commit crimes in specific target areas."[315] In order to rank chronic offenders, individuals were assigned "points" based on factors like gang membership and their history of interactions with the police. Those who reached a specific threshold of points were placed on the Bulletin, which had a two-step identification process, according to an investigation conducted by *The Intercept*:

> An initial screening phase, in which a "crime intelligence analyst" subjectively decides whether the police records, like arrest reports and field interview cards, associated with an individual are "relevant" enough to move them to a "workup" phase. The "workup" involves software provided by Palantir that pulls data on criminal history and affiliations, and from license plate readers and social media networks, and uses it to create a "chronic offender score" for the individual. Once someone is deemed a sufficient threat based off their score, officers send them letters and are encouraged to knock on their doors to let them know they're being monitored. Officers are also instructed to look out for opportunities to stop or arrest them (if they have a warrant out).[316]

Under Operation LASER, each new point of contact with police earns a person one additional point.[317] For Brayne, this system results in a self-perpetuating cycle, in which individuals in historically overpoliced neighborhoods were "more likely to be stopped, thus increasing their point value," justifying their increased surveillance, and making it more likely that they will be stopped again in the future.[318] As Brayne writes:

> Despite the stated intent of the point system to avoid legally contestable bias in police practices, it hides both intentional and unintentional bias in policing and creates a self-perpetuating cycle: if individuals have a high point value, they are under heightened surveillance and therefore have a greater likelihood of being stopped, further increasing their point value. Such practices hinder the ability of individuals already in the criminal legal system from being further drawn into the surveillance net, while obscuring the role of enforcement in shaping risk scores. Moreover, individuals living in low-income, minority areas have a higher probability of their "risk" being quantified than those in more advantaged neighborhoods where the police are not conducting point-driven surveillance.[319]

According to activist Jamie Garcia, these LASER zones helped the LAPD statistically compute and compile arrest data, crime reports, and other information related to gun violence. Garcia was one of several activists who sued the city of Los Angeles in 2018 on behalf of the Stop LAPD Spying Coalition, alleging officials failed to comply with a public records request seeking more information about Operation LASER:

> The police will essentially go to these LASER zones, but being in those LASER zones, they look for particular people. [...] They look at [field interview] cards. They look at full compliance units. They look at arrest data. They look at calls for service, and they essentially have a criminal intelligence detail.[320]

Such practices hinder the ability of individuals already in the criminal legal system from being further drawn into the surveillance net, while obscuring the role of enforcement in shaping risk scores. Moreover, individuals living in low-income, minority areas have a higher probability of their "risk" being quantified than those in more advantaged neighborhoods where the police are not conducting point-driven surveillance.

Only one study, authored by the designers and creators of the program, has evaluated the efficacy and success of the LASER program. Unsurprisingly, the study found that the use of LASER correlated with a reduction in gun violence.[321] Though the LAPD instructed officers to keep track of their "dosage," the frequency with which a squad car drove through each designated LASER zone, an audit released in 2019 found that dosages were inconsistently monitored.[322]

The program was terminated by the LAPD in 2019 after "members of the department's civilian oversight panel questioned the effectiveness of data-driven strategies that rely on algorithms and other computer technology to identify violent offenders and map out areas most prone to criminal activity."[323] A report from the office of LAPD Inspector General Mark Smith, published in 2019, additionally found that the data-driven policing strategies employed by the LAPD "lacked oversight and that officers used inconsistent criteria to label people who were likely to commit violent crimes,"[324] resulting in insufficient data to accurately measure the programs' success. These findings reflect those of programs in other cities, like the SSL, that were subject to similar scrutiny.

4. New York: Operation Crew Cut

The New York City Police Department (NYPD) maintains an expansive and growing database of individuals it claims are "gang-affiliated."[325] The database has included the data of as many as 42,000 people, ninety-nine percent of whom are Black and Latinx and none of whom can challenge their inclusion in the database.[326] In a growing number of cases, the data that populates these databases also come directly from social media platforms, which have dramatically expanded the scope of law enforcement surveillance in intelligence-gathering, data collection, and criminal investigations. The use of social media surveillance to populate the gang database functionally creates "a self-fulfilling prophecy in which basic social media etiquette is mistaken for membership in a criminal enterprise."[327]

As early as 2012, the NYPD secretly contracted with Palantir to analyze the data it collects.[328] Neither the public nor the City Council were aware of the NYPD's use of Palantir until five years later, when the pair got into a contract dispute.[329] As a result, little is publicly known about what kinds of data the NYPD fed through Palantir. But the Manhattan District Attorney's office developed a similar gang database (the "Crime Prevention System") that is still supported by Palantir.[330] The NYPD currently operates its own analytical system, known as "Cobalt," which the department developed in-house in partnership with IBM.[331]

"Operation Crew Cut" emerged from the growing prevalence of social media, the NYPD's long history of targeting people it purports to be members of street gangs, and mounting law enforcement surveillance.[332] At the end of 2012, New York City Police Commissioner Ray Kelly announced the launch of Operation Crew Cut as a new anti-gang initiative that would target "looser associations of younger men who identify themselves by the block they live on, or on which side of a housing development they reside."[333] Kelly explicitly tied Operation Crew Cut to the use of social media monitoring and the rising impact of technology on police missions.

Targeting individuals as young as 10 years old through online surveillance, Operation Crew Cut demonstrated that the NYPD had no age limits on their protocols for social media investigation. Adult officers routinely watched alleged crew members by posing as young people, "[with] no official provision requiring parents or guardians to be notified."[334] In his meeting with the Task Force in 2019, Jarrell Daniels, a research assistant at the Center for Justice at Columbia University who was incarcerated for four years after an Operation Crew Cut sweep, stated that the primary evidence used for his indictment were posts on his Facebook account. Daniels was one of ten people, all Black and Latinx men and predominantly 18 and 19 years of age, arrested on charges, including conspiracy to commit murder, assault, weapons possession, and drug possession and sales.

In an investigation of Operation Crew Cut's aftermath published by *The Appeal*, Daniels found that he was left in the dark about the evidence being used against him in his case due to New York's restrictive laws on pretrial criminal discovery. According to Daniels, it was not until a *New York Post* article[335] published the day after his arrest that he learned his Facebook account was being used against him. As Daniels described in his interview with the Task Force:

> We didn't know that we were being investigated at all. And that's kind of how they want the policy to function, is that they don't want you to know you're being investigated. It's not a physical investigation, so you won't see unmarked cars sitting in a neighborhood. You won't see people, like, doing stakeouts, the way policing used to happen. They sit behind a desk and navigate social media as if they're a young person.[336]

Years before Daniels was arrested under Operation Crew Cut, the criminal legal system had already "cast a long shadow on him."[337] During his first week of high school, Daniels was arrested for gang assault and robbery after a fight broke out several blocks from a bus he was riding. Though the charges were dropped three years later, the "damage was already done"[338]:

> My self-image was being seen through the systems I couldn't escape. I was no longer a son, a brother, or even a student. Quite frankly, I wasn't even a teenager anymore. I was a soldier whose sole purpose was to survive. Thinking about anything beyond that would just cause me pain.[339]

In an interview with *The Guardian* in 2015, defense attorney Andrew Laufer asserted that social media-driven initiatives like Operation Crew Cut were leading to an uptick in arrests on mass conspiracy charges. Laufer referred to these strategies as "round-ups" that specifically target low-income minority communities with the objective of meeting high quotas of fines or arrests.[340] The most prominent example of such a round-up was the "Bronx 120 raid," in which 120 people, almost all young Black and Latinx men, were indicted during what prosecutors have called the "largest gang takedown in New York City history."[341] To secure these indictments, "prosecutors relied heavily on text messages and social media posts to prove connections between defendants, essentially turning friendships and social relationships into evidence of conspiracy."[342]

According to Babe Howell, a legal scholar and gang policing specialist at City University of New York School of Law, such strategies employed by the NYPD were ultimately an expansion of the stop-and-frisk regime, with no effective means of oversight to limit the extent of surveillance or information being collected:

> The intensive surveillance extends to following Twitter feeds, monitoring Facebook (often by creating fake profiles of attractive young women), and monitoring YouTube videos. Whether the police should be engaged in this level of surveillance of youth for intelligence collection purposes, without any prior showing or justification, is an important question that merits serious consideration and is not one that should be answered in a kneejerk manner based on our fear of gangs. Police lists may be shared with immigration or potential employers and cause substantial collateral damages even in the absence of criminal convictions or arrests.[343]

The Bronx 120 raid was not the first of its kind. In 2014, a raid in Harlem, which at the time was also billed as "the largest" in New York's history, led to 103 indictments. These raids follow a troubling history of the NYPD rapidly expanding their gang databases by targeting individuals using dubious and nontransparent allegations of gang membership.[344] In his meeting with the Task Force in 2019, scholar and activist Josmar Trujillo elaborated on how gang policing initiatives like Operation Crew Cut simply serve to criminalize young people based on their social networks:

> They are adding tremendous amounts of people to the database with no transparency or process around it. This is happening at the same time that we are seeing more and more gang raids happening. All of this is connected, what information they are pulling in. And the only way that I can reason in my head that they are adding more peoples to the database, is that they are making assumptions that associations using big data to lump people in and throw them in to the database.[345]

Like the NYPD, police departments throughout the country have constructed gang databases, often localized within cities or similar jurisdictions. These databases assemble and maintain a broad swath of identifying data on individuals known or suspected to be gang members, or associated with gang members. While police departments have articulated that these databases are needed to reduce gang-related criminal activity,[346] individuals included in a gang database are subjected to increased police surveillance and monitoring, and can also face enhanced criminal charges upon arrest. In addition, prosecutors cite to their placement on these databases during bail arguments. Once convicted, people in gang databases can receive longer prison sentences through gang enhancement statutes.[347] For non-citizens, inclusion on a database can even lead to ICE detention and deportation.[348]

Placement on a database can even carry the weight of collateral consequences, which are usually tied to convictions, as information on these databases "has been shared with employers, landlords, public housing, and school administrators, often leading to additional punishments, evictions, and exclusion from services and resources."[349] Despite claims that the NYPD "routinely expunged names from the database to eliminate people who are no longer affiliated with these groups,"[350] an investigation in 2019 revealed that the NYPD, does in fact, routinely disseminate sealed data to third parties, including prosecutors, the news media and housing, immigration and family-court officials.[351]

Under two New York state laws from 1976[352] and 1980,[353] law enforcement is only supposed to be able to access sealed records with a court order, or when conducting a background check on someone under narrow exceptions. Even in states like New York with statutes explicitly protecting sealed arrest records, "the burden is often on individuals themselves to hire a lawyer, go to court, pay a fee, and prove they should have their criminal history — which, again, never resulted in any conviction — expunged. This can be additionally difficult because, for the very reason that such arrests often did not result in charges being filed, there may not be any court records to get expunged, even if the police still have the information on their own computers."[354]

CRITICAL ANALYSIS OF DATA-DRIVEN POLICING

Though it is tempting to view the recent explosion of data-driven policing as another inevitable consequence of the big data revolution, an inordinate number of resources have, in fact, been invested by criminal justice bureaucrats, police departments, and technology companies in designing, developing, and deploying these technologies.[355] With machine learning algorithms now playing a role in every aspect of policing – including where to deploy officers, whom to surveil, and whom to encounter – such technologies can no longer be divorced from the explicit and implicit biases that police officers bring to their patrols.

A growing body of research and journalism has shown that use of predictive algorithms in policing– which primarily use and are trained on historical crime data – replicate and amplify existing systemic biases, often with little to no thought given to how "different crime-reduction policies, crime legislation, profiling tendencies, or sentencing biases influence the patterns found by [such] algorithms in the data."[356] Intensified public scrutiny of these algorithms has additionally raised questions about how they are developed, implemented, and marketed; why they are not subject to more review; and whether there are mechanisms in place to properly assess their risks, vulnerabilities, and potential for greater societal harm.

To form its recommendations for this report, the Task Force met with a diverse group of witnesses, including technologists and industry experts, law enforcement personnel, academics, attorneys, advocates, and community stakeholders. They also read extensive literature on data-driven policing technologies, including but not limited to, books, academic articles, studies, reports, and news articles. The Task Force arrived upon the over-arching recommendation that data-driven policing technologies are ineffective; lack scientific validity; create, replicate and exacerbate "self-perpetuating cycles of bias";[357] and hyper-criminalize individuals, families, and communities of color. These technologies are moving the criminal legal system from mass incarceration to mass criminalization.

This section of the Report will address the Task Force's recommendation, and interrogate the significant information gap separating the developers who create and design technology; the police departments that utilize the technology; the communities where the officers utilizing, relying on, and reacting to the technology are deployed; the individuals who are charged with crimes connected to the technology; the defense attorneys who represent those individuals; civil liberties attorneys; and the courts.

A. METHODOLOGICAL PROBLEMS

Though prediction has always been a fundamental part of policing,[358] the emergence of predictive algorithms in policing was considered particularly novel for its alleged ability to apply artificial intelligence to quantities of data once considered too large in quantity and too complex for police departments to analyze.[359] Its proponents have since claimed that predictive policing programs can lower crime, revolutionize public safety, and help under-resourced departments "do more with less,"[360] while critics have argued that such programs produce self-perpetuating feedback loops of crime prediction, placing historically over-policed individuals and communities at further risk of harm. The Task Force found the latter, that these programs entrench existing

biases and exacerbate the disproportionate impact of policing on BIPOC, low income and other marginalized communities.

1. Garbage In, Gospel Out

Though big data tools may appear to provide an "objective analysis"[361] of information, the basic building blocks[362] of a predictive software program involve many human discretionary decisions, beginning from which mathematical model to use to the geographic areas in which the tool is eventually implemented. Algorithmic policing tools may vary in the specific types of information they employ, but they all ultimately rely upon historical crime data compiled by the police themselves.[363]

Though many may assume that police data is objective, it is embedded with political, social, and other biases. Indeed, police data is a reflection of the department's practices and priorities; local, state or federal interests; and institutional and individual biases. In fact, even calling this information 'data' could be considered a misnomer, since "data" implies some type of consistent scientific measurement or approach. In reality, there are no standardized procedures or methods for the collection, evaluation, and use of information captured during the course of law enforcement activities, and police practices are fundamentally disconnected from democratic controls, such as transparency and oversight.

Studies by criminologists, legal scholars, and technologists have repeatedly shown that historical crime data is at best a "partial representation of crime in the community,"[364] and at worst, a record containing "falsified crimes, planted evidence, [and] racially biased arrests."[365] Many have since proposed that such crime statistics should instead be regarded as a record of law enforcement's "response"[366] to what actually happens in a community, since police officers "wield significant amounts of discretion about what they see and know about crime."[367]

For example, in 2019, legal scholar Rashida Richardson identified nine jurisdictions where predictive policing systems were trained on police data generated during periods when the department was found to have engaged in various forms of unlawful and biased police practices[368]:

> Though many may assume that police data is objective, it is embedded with political, social, and other biases. Indeed, police data is a reflection of the department's practices and priorities; local, state or federal interests; and institutional and individual biases. In fact, even calling this information 'data' could be considered a misnomer, since "data" implies some type of consistent scientific measurement or approach. In reality, there are no standardized procedures or methods for the collection, evaluation, and use of information captured during the course of law enforcement activities, and police practices are fundamentally disconnected from democratic controls, such as transparency and oversight.[369]

As legal scholar Elizabeth Joh noted in her conversation with the Task Force, the discussion surrounding big data policing programs often assumes that the police are the consumers, or the "end users," of big data, when they themselves are generating much of the information upon which big data programs rely from the start.[370] Prior to being fed into a predictive policing algorithm, crime data must first be "observed, noticed, acted upon, collected, categorized, and recorded"[371] by the police. Therefore, "every action – or refusal to act – on the part of a police officer, and every similar decision made by a police department, is also a decision about how and whether to generate data"[372]:

> Even more reflective of police discretion are field contact cards: information officers collect about people they encounter on the street for consensual, information-producing conversations. Contact cards are unlikely to have an even or random distribution. Once transformed into data, this information can appear neutral and objective, even though they are the products of individual discretionary decisions. Moreover, these highly discretionary decisions can be further influenced by other ones, such as departmental pressures to 'produce' contact cards, or by metrics that assess officer productivity through consensual contacts, stops, and arrests.[373]

If crime data is to be understood as a "by-product of police activity,"[374] then any predictive algorithms trained on this data would be predicting future policing, not future crime;[375] as statisticians Kristian Lum and William Isaacs put it, this functions as a self-fulfilling prophecy. Neighborhoods that have been disproportionately targeted by law enforcement in the past will be overrepresented in a crime dataset, and officers will become increasingly likely to patrol these same areas in order to "observe new criminal acts that confirm their prior beliefs regarding the distributions of criminal activity." [376] As the algorithm becomes increasingly confident that these locations are most likely to experience further criminal activity,[377] the volume of arrests in these areas will continue to rise, fueling a never-ending cycle of distorted enforcement.[378]

The biases held by police officers and those reporting crimes, and correlations between attributes like race and arrest rates, will not only be recognized and replicated by the algorithm, but directly integrated into the software "in a way that is subtle, unintentional, and difficult to correct, because it is often not the result of an active choice by the programmer."[379]

Lum and Isaac conducted a pivotal study in 2016[380] that exemplified the biases inherent in historical crime data and the racial disparities in predictive policing using two sets of data for comparison: estimates from a public health survey about patterns of illegal drug use, and predictions based on the PredPol algorithm using data from the Oakland Police Department's record of drug crimes from 2010.[381]

According to Lum and Isaac, "a comparison of these figures [told] dramatically different stories about the pattern of drug use"[382] in the city of Oakland, with the PredPol model consistently targeting lower-income historically Black neighborhoods that were "already over-represented in the historical police data."[383] Though data from the National Survey on Drug Use and Health had suggested that drug use was roughly equivalent across all racial classifications, by plugging historical crime data into the PredPol model, Black people were targeted by predictive policing at roughly twice the rate of white people.[384]

While developers of predictive policing tools have claimed that data reported by victims of crime may be able to capture a more accurate picture of crime rates, researchers Nil-Jana Akpinar and Alexandra Chouldechova found that the predictions obtained using victim report data were also skewed.[385] After training the PredPol model on victim report data from Bogotá, Colombia, one of very few cities for which independent crime reporting data is available at a district-by-district level, they found that the tool still made significant errors, repeatedly predicting high rates of crime in areas where few crimes had been previously reported.

Akpinar and Chouldechova demonstrated that differential victim crime reporting rates across geographical areas can lead to outcome disparities in common crime hot spot prediction models even when adjusting for disparities in the data, leaving potential for both over-policing and under-policing certain neighborhoods. Simply put, predictive policing systems are only as good as the data that they possess.[386] Likewise, programs like Palantir's Gotham platform contain "so much data from so many sources"[387] that there is significant potential for inaccuracy and misuse. Palantir has the alleged ability to "intake suspicious activity reports from across the many law enforcement agencies in the region, compare them against each other and all sources of intel…and identify links or patterns of suspicious behavior."[388]

For example, in states like California, which once accounted for close to 90% of the sales of Palantir's systems to domestic law enforcement,[389] Palantir has integrated various law enforcement and governmental databases into its system, ranging from criminal records and restraining orders, to the details of cars and drivers.[390] Adding this much data to Palantir's system "expanded the potential for misusing that information," and with key analyses now being entirely automated, such algorithmic filtering could "produce results of wildly variable quality."[391]

These issues of biased data are well-known to the companies that design and sell data-driven policing technologies. As Lum and Isaac lay out, the barriers to correcting for this degree of bias are both largely unaddressed and nearly insurmountable:

> While several prominent predictive policing vendors have acknowledged concerns about the inclusion of biased data in their systems, most vendors fail to account for these structural and systemic errors

> in the data, often overestimating what can be remedied. Not only is the challenge of identifying and correcting these problems difficult, if not insurmountable, but it also raises significant doubts about the ability to distinguish known problematic data categories, such as drug-related arrest data, from data categories that are customarily considered objective, such as calls for service data. Moreover, even where such distinctions are possible, they would have to occur on a jurisdiction-by-jurisdiction basis, since police data collection and classification practices vary by department and are often performed in ways that make aggregate or comparative analysis impossible.[392]

With an increasing number of police departments already succumbing to the "pressures of managerial techniques that emphasize quantitative measures of effective policing,"[393] some experts have suggested that data-driven policing strategies and tools have facilitated the return of broken windows policing. Under such a logic, officers "are explicitly encouraged to look for and harshly penalize petty crime that may go unnoticed in other neighborhoods."[394] In her meeting with the Task Force, criminal defense attorney Michelle Fields described predictive policing as "stop-and-frisk" but with numbers and "bad data."[395]

Meanwhile, Vincent Southerland, Executive Director of the Center on Race, Inequality, and the Law at NYU School of Law, noted to the Task Force that predictive policing systems were inextricably tied to the same patterns of over-policing and hyper-criminalization that characterized broken windows policing in the 1980s:

> In those instances [where police departments] were using the PredPol tool, what you have is essentially a police officer being told, "Look, a crime is going to happen in this particular community. Be on the lookout for crime": almost priming them to engage in racial profiling and in this heightened level of suspicion of individuals who are walking around…priming police to engage in this misconduct based on their own [past] interactions with the community.[396]

This leads to the question, as Andrew Ferguson puts it: "How do you fix an error in the data if you cannot see that such an error exists?"[397] Contrary to popular belief, no predictive policing software is self-executing, and "an algorithm that relies on data produced by biased institutions and attitudes does nothing to inherently remove that institutional bias."[398] Research and scholarship have repeatedly shown that crime data cannot reflect the rate at which crimes are committed; crime data can only reflect the rate at which crime was "caught and recorded"[399] by the police. This raises the questions of whether "crime data" can be considered reliable as data. Indeed, all predictive policing systems run the risk of being built on an incomplete and biased understanding of where crimes take place and who is actually committing them,[400] with real costs to communities already under pervasive police scrutiny and surveillance.

2. Façade of Algorithmic Transparency

The 2001 film *The Minority Report* opens with a montage of visions from "pre-cogs": mutants who can predict the future and help the police stop crime before it happens. Since the emergence of big data tools in policing, technology companies have frequently touted their predictive data-driven policing programs as almost *Minority Report*-style technologies with the purported ability to "predict the future." After all, as legal scholar Andrew Ferguson writes, "a black-box futuristic answer is a lot easier than trying to address generations of economic and social neglect, gang violence, and a large-scale underfunding of the educational system."[401]

Today's officers are not only collecting massive amounts of information about individuals,[402] but are also identifying "suspicious" persons and activities on an unprecedented scale.[403] Like the "pre-cogs" in Philip Dick's dystopian universe, it has been well-documented that predictive policing systems, and data analytics more broadly, confer a general perception of "empirical neutrality and infallibility";[404] users of such systems frequently overestimate the accuracy, objectivity, and reliability of information that comes from a computer program.[405] As research has repeatedly demonstrated, however, an "algorithm is only as infallible as the human beings who choose the variables, input the data, and act on the results."[406] Technology and civil liberties attorney Matt Cagle further elaborated on this point in his meeting with the Task Force:

> Technologies are not operating in a vacuum. They're operating in real communities. They are ingesting certain types of data about certain parts of communities more than other types of communities, and they're feeding on a history of policing patterns[...] [A] history of what gets prioritized in terms of enforcement, at a particular department, may also influence the data that goes into the algorithm that then dictates what sort of decisions are made on the other side of the black box.[407]

Since people have a tendency to believe a computer-generated report over that of a human-created report, predictive policing programs and other automated decision-making systems often run the risk of "being trusted above human judgment while simultaneously concealing potential unchecked errors."[408] Biases in machine learning algorithms pose a "particularly insidious risk to disadvantaged groups by creating a pseudo-scientific justification for discriminatory treatment,"[409] While transparency can help prevent deliberate or semi-deliberate discrimination, it cannot singlehandedly "correct the effects of the unintentional, institutional discrimination embedded in the data itself."[410] As legal scholar Lindsey Barrett writes, this is arguably the most serious and difficult concern to counteract:

> When the cause of a flawed result produced by a machine-learning algorithm is unknowable, transparency will do little to solve the underlying problem, apart from the value of revealing that the problem exists. Greater transparency is also unlikely to correct flaws in application, such as automation bias, and, to the extent that it leads to better programs, it can only aid in preventing arbitrary or discriminatory policing. If the use of these algorithms is transparent, but does not lead to the correction of encoded bias in the data or the use of poor-quality information, transparency is fairly hollow as an institutional principle.[411]

3. Problematic Predictions in Practice

In a stated effort to demonstrate transparency and uphold the privacy and civil rights of its users,[412] predictive policing companies like PredPol have previously published the algorithms and methodologies they employ in their software. According to academic papers co-published by PredPol's founders, including Jeff Brantingham, George Mohler, and Sean Malinowski, criminal activity and seismic activity supposedly follow "surprisingly similar patterns."[413] These materials, which are frequently cited in PredPol materials, are often referenced as proof that greater crime reductions could be achieved by improving predictive algorithms.

Technologies are not operating in a vacuum. They're operating in real communities. They are ingesting certain types of data about certain parts of communities more than other types of communities, and they're feeding on a history of policing patterns[...] [A] history of what gets prioritized in terms of enforcement, at a particular department, may also influence the data that goes into the algorithm that then dictates what sort of decisions are made on the other side of the black box.

In his meeting with the Task Force, Philip Stark, Associate Dean of Mathematical and Physical Sciences at the University of California, Berkeley, pushed back against PredPol's findings in the aforementioned study, stating that their use of seemingly "complicated-looking mathematical formulas"[414] like an earthquake prediction model intentionally "gives an air of objectivity"[415] to an algorithm that is fundamentally flawed and not scientifically proven to work as they may claim. Though the earthquake prediction model is widely employed to predict and describe the occurrence of earthquakes, there has been little discussion dedicated to the limits of, and influences on, its estimations, even in seismology.[416] According to Stark:

> If [the earthquake prediction model] does not work where it was invented in situations where the input data are essentially perfect, and moreover if [an] algorithm that only takes a few lines of code to implement does just as well in earthquakes, then why would we expect [it] to work well in a policing situation where the data aren't complete, aren't accurate? [...] What they're predicting is not crime, it's what happens if you send police there.[417]

For Stark, PredPol's program deliberately exploits the fact that specific kinds of crime have a tendency to "cluster" together in space and time, embedding hidden assumptions and policy decisions that may not be visible to those developing and using these algorithms. Sociologist Alex Vitale perhaps put it best in his book *The End of Policing*: the problem is not police training, police diversity, or police methods; rather, "the problem is policing itself."[418] Proponents of data-driven policing are typically so focused on optimizing existing practices that they are unable to answer, or even ask, questions about what should be done with the predictions that are made. As mathematician Ben Green writes:

In the hands of police, even algorithms intended for unbiased and nonpunitive purposes are likely to be warped or abused. For whatever its underlying capabilities, every technology is shaped by the people and institutions that wield it. Unless cities alter the police's core functions and values, use by police of even the most fair and accurate algorithms is likely to enhance discriminatory and unjust outcomes.

> In the hands of police, even algorithms intended for unbiased and nonpunitive purposes are likely to be warped or abused. For whatever its underlying capabilities, every technology is shaped by the people and institutions that wield it. Unless cities alter the police's core functions and values, use by police of even the most fair and accurate algorithms is likely to enhance discriminatory and unjust outcomes.[419]

In Chicago, for instance, the Strategic Subjects List (SSL) program, which was originally conceived to address the spread of gun violence in social networks, was later converted into a pervasive tool for surveillance and over-criminalization. Although the original stated intention for the SSL was to address gun violence as a public health issue, it largely ended up being used as a surveillance tool that disproportionately targeted people of color,[420] and consequentially led to increased arrests of predominantly Black and Brown young men. This was confirmed by a 2016 RAND Corporation study, during which RAND researchers were "allowed to view the list, sit in on internal meetings, and generally observe how the tool was being used." The researchers learned that "CPD wasn't using the list as a way to provide social services; instead, CPD was using it as a way to target people for arrest."[421]

Unlike the SSL, independent empirical studies have yet to be conducted on Palantir Technologies' highly secretive data-driven policing systems. Law enforcement may account for a small portion of Palantir's business, but the departments that deploy Palantir "are also dependent upon it for some of their most sensitive work, [...] spotting links and sharing data to make or break cases."[422] Palantir's Gotham platform, which is primarily marketed to law enforcement, can allegedly ingest and sift through millions of digital records across multiple jurisdictions. However, the company doesn't disclose the full variety of data that go into the system, nor the algorithms used to create and track profiles of individuals.[423]

By joining Palantir's coveted and extensive data-sharing network, "customers must rely on software that only the company itself can secure, upgrade, and maintain"[424] because the software and support services are considered "proprietary" to the company. Additionally, sensitive government data becomes privatized when law enforcement departments enter into contracts with Palantir and process data using their system. In her meeting with the Task Force, sociologist Sarah Brayne elaborated on this point of contention, as it pertained to her fieldwork with the Los Angeles Police Department (LAPD):

> Palantir goes to great lengths to emphasize that they are just this like infrastructure that sits on top of the data. [The] NYPD tried to cancel their contract with Palantir, but investigators have been building up all of these different networks and places and criminal histories, [so that] when they terminate their contract with Palantir, they don't get all of that linked data. They just get the raw input data from like five years ago, losing five years' worth of intelligence.[425]

An investigation published by *Wired* in 2017 subsequently revealed a history of police departments accusing the company of "spiraling prices, hard-to-use software, opaque terms of service, and 'failure to deliver products.'"[426] Public contracts and other records obtained by *Wired* indicated that new users are initially "welcomed with

discounted hardware and federal grants," while also being incentivized to share "their own data in return for access to others'"; consequentially expanding Palantir's universe of databases:

> When enough jurisdictions join Palantir's interconnected web of police departments, government agencies, and databases, the resulting data trove resembles a pay-to-access social network — a Facebook of crime that's both invisible and largely unaccountable to the citizens whose behavior it tracks.[427]

For example, after the NYPD cancelled its contract with Palantir in 2017 and requested copies of Palantir's analyses, the company declined to provide it in a standardized format that would work with the NYPD's new system; saying that doing so, would expose its intellectual property. The standoff highlighted a recurring issue for companies and governments that outsource their data-mining tasks to outside contractors like Palantir.[428]

B. TRANSPARENCY, TRADE SECRETS, AND NON-DISCLOSURE AGREEMENTS

Meaningful transparency, as detailed in the last section, is more than making an undecipherable algorithm public. It gets to the heart of what defense lawyers, local legislatures, and communities know and understand about the tools being used to police the public.

> As Jay Stanley, a Senior Policy Analyst for the ACLU, told the Task Force, transparency is important, number one, because these things are very brand-new, and they're also very difficult. In many ways you need the thousand eyeballs. You need experts, academic experts, hackers, people in the community to able to see what the police are doing so it can be evaluated properly because if you're just getting the company that's offering the product telling you that it's working, it's not worth a lot. [...] The degree to which we want our police mucking around in our social media or gathering all kinds of data about us and crunching through numbers is a decision that we need to make democratically. And we can't even make those decisions if we don't know what's going on.[429]

Lack of transparency ultimately exists at every level of a predictive policing system.[430] As with most surveillance technologies, the process by which they are procured and the software that they run are hidden from defense attorneys, those accused in criminal cases, the general public, and at times even local elected officials. The lack of transparency jeopardizes fundamental constitutional rights and prevents people from understanding how their communities are policed.

Ultimately, "there is no way for law enforcement, courts, legislatures, or the public to gauge the accuracy and value of [a] software without understanding how the methodology led to any purported success."[431] For this reason, predictive policing algorithms are often referred to as a "black box" because the "calculations used to make a decision may be inscrutable to the person affected by that decision."[432] This inscrutability prompts "calls to require legal rights for individuals to know the basis of automatic decision-making affecting them,"[433] and for these systems to be made auditable by the public.

1. Private Companies: "Trade Secrets," Non-disclosure Agreements, and Profit

Companies that build data-driven policing technologies claim proprietary rights over their methodologies. They have asserted such claims in response to subpoenas seeking information about data-driven policing technology (and, really, any other technology made to inform and assist law enforcement). According to Rebecca Wexler, a Professor at University of California, Berkeley School of Law and the Faculty Co-Director of the Berkeley Center for Law and Technology, in her meeting with the Task Force, companies make "trade secret" claims to protect these critical details from disclosure. Thus, companies "claim entitlements to withhold th[e] information from defense attorneys and those accused in criminal cases, refusing to comply even with... subpoenas that seek information under a protective order and seal." She explained that "[t]he current state of trade secret law enables those agencies to claim proprietary protections."[434]

As Wexler stressed to the Task Force, "[t]he automation of criminal justice decision- making is creating these conflicts between developers' intellectual property rights and defendants' access to evidence."[435] Legal scholar Elizabeth Joh has conveyed some of the ways in which these conflicts have exacted harm on defendants:

> When new surveillance technologies are kept secret because of non-disclosure agreements, they cannot be challenged by criminal defendants and these challenges can't be decided by judges whatever the merits of the defendants' claims. The use of a new surveillance technology may or may not be considered a Fourth Amendment search, but a private company's insistence on secrecy removes the legal issue from judicial review.[436]

As is obvious from claims of trade secrets and the conditions cemented in non-disclosure agreements, the companies that produce data-driven policing technology face competitors, known and unknown. These companies are in the *business* of selling and profiting from the data-driving policing technology they produce.

> [E]ven without explicit nondisclosure agreements, big data tools can remain secret because they contain proprietary information that companies may be unwilling to release. Nor are private companies producing these tools subject to public records laws that would require them to divulge relevant and useful information.[437]

In his meeting with the Task Force, Daniel Kahn Gillmor, Senior Staff Technologist for the ACLU's Speech, Privacy, and Technology Project also suggested that the business of producing these algorithmic systems in policing has given rise to a new "surveillance economy," in which the financial prospects of an increasing number of companies and private foundations are "based on the amount of information they can collect on people."[438] Though the public may or may not have municipal control over what the police do, Gillmor argued that it was even more unlikely that the public could obtain "municipal control over what every corporation that gathers information about people does."[439] As Gillmor elaborated:

> We need these systems to be auditable by the public, [but] many of the proprietary vendors are uninterested in being auditable... I think it's worth asking someone who's proposing a system like this for details of what specifically — what data is being drawn from, how the data is being combined, and where the data will be used to drive policy.[440]

The lack of transparency surrounding data-driven policing and the industry's fundamental business model are interwoven. Private companies compete to build, market, and sell the technology. Police departments, therefore, "are customers or clients of private companies"[441] — some companies even provide technology to police departments initially for free, with the goal of selling departments on the need to continue using the technology. As Rashida Richardson, a visiting scholar at Rutgers Law School and the Rutgers Institute for Information Policy and the Law, explained to the Task Force, many departments themselves often possess little to no insight into the inner workings of the systems they employ and lack incentive to do so without explicit transparency measures in place:

> [A] more public and transparent and accountable process would also just look bad for [police departments] everywhere, when they are kind of like, "I don't actually know how this system works. We are just using it because we got it for free, and we thought it would help." When you don't have data to prove its efficacy as backup, which they often don't, then it raises lot of good governance questions: like why are you using these systems, why are public tax dollars being used for this?[442]

In addition to paying the companies to continue using the technology after the end of their free trial period, some police departments also pay hefty subscription fees to these companies to upload police-generated data.[443] For example, documents obtained by a public records request from *Wired* in 2017 revealed that police departments in Los Angeles were paying around "$122,000 each" for over one hundred Palantir servers in order to "maintain intelligence data processing capacity and capability,"[444] with the expectation that they would purchase more as they continued to store and process data using Palantir's software. Accordingly, the companies have a financial stake in building police dependence on the technology, as well as the menu of other related services.

2. Proprietary Algorithms and Frustrated Judicial Processes

Companies hide behind a number of legal tactics to protect their technologies from independent scrutiny. They argue that revealing the software, training data, or other details underlying their product would damage

their ability to develop and market their products. Critics believe that independent scrutiny would reveal the infirmities of their products, thereby harming companies' bottom line. In either case, the information remains shielded from the public.

> The use of nondisclosure agreements, the ability to dominate a particular market, and the shielding of proprietary information all share a common feature: They exert an undue influence by private companies on public police practices. That influence can and has resulted in real harms that affect legal change, police oversight, and police accountability.[445]

Often, defense lawyers are not aware of or do not have access to information about the surveillance tools that were behind their client's arrest and prosecution. The lack of information is part of a larger trend in privatization in the criminal legal system.[446] The end result is that private companies have great deal of influence over the indicators that will determine who is stopped, searched, and arrested, to the detriment, legal and otherwise, of the accused. Elizabeth Joh lays out the problem:

> Police that rely on big data tools to identify those people and places that deserve attention are using these programs to help develop their own assessments about suspicion. These assessments in turn can help develop the legal suspicion necessary to conduct stops, frisks, and arrests. At some point in the near future, courts will have to determine whether an algorithm's determination can form the basis, at least in part, of Fourth Amendment suspicion.
>
> If informants and tips can help develop reasonable suspicion, it is likely that courts will accept big data analysis as another source of information for the police as well. The problem for courts and defendants hoping to find out how a big data program has arrived at its conclusions is that the suspicion itself has been outsourced, at least in part. How an algorithm recommended police attention to one person or city block rather than another may be guarded as a "trade secret" that the algorithm's creators are unwilling to reveal.[447]

Thus, by outsourcing the elements of suspicion and then hiding those determinations between trade secret claims, the judicial process is frustrated in its ability to assess the constitutionality of law enforcement actions.

3. Communities, Local Governments, and Police Departments

One of the overarching concerns reiterated to the Task Force about data-driven policing technology is the lack of notice and information provided to impacted communities related to these tools. The relationship between the private companies that build these tools and police departments is insular and exclusive. With exceptions in a very small number of jurisdictions, police departments do not inform impacted communities of their desire to deploy these tools, do not provide a justification to the impacted communities for the tools, and do not disclose the actual use of these tools or the policies that govern their use (to the extent that there are any).

In turn, residents of impacted communities do not have the opportunity to learn about the technology, offer input to police departments on proposals to deploy the technology, or provide any oversight of the police department's use of the technology once it has been deployed. Rather, residents are left in the dark. As Josmar Trujillo told the Task Force, "We've…never been asked about what we think about [predictive policing] now that it's going forward. There's been no input. There's no side of predictive policing that is community based."[448] As a result, data-driven policing technology exacerbates both the hyper-criminalization and hyper-marginalization of these impacted communities. Residents have no voice in matters that impact their lives disproportionately.

The lack of transparency also undermines democratic governance. Not only are residents denied any opportunity to be informed about and involved in the process of determining whether police departments should be permitted to deploy data-driven technology in their communities, they are also shut out of considering, and providing input on, critical questions that are central to accountability, transparency, and democratic governance. One of these questions is: who will (or should) pay for these data-driven tools? Another, perhaps more fundamental question is: what exactly do these tools do in and to the communities where they are deployed? As Chad Marlow, senior advocacy and policy counsel at the ACLU, told the Task Force:

> Right now, we have the problem of not even knowing if this technology is being used, let alone having the democratic principal of having the people and their elected officials have a voice in deciding whether the benefits of using these technologies outweigh the risks, not just the benefits to law enforcement, but the costs, and whether the risk can be appropriately reigned in.[449]

Recently, several communities have learned that law enforcement deployed data-driven technologies in the places where they reside. In response, community stakeholders have been working at the local and state level on advocacy efforts to pass laws and ordinances that would enforce democratic governance on data-driven policing technology. Such laws often require the disclosure of any and all surveillance technologies that law enforcement is considering for use — or is actually using, or has used — and the creation of systems to give community members and legislators meaningful opportunities to weigh in on whether the technology should be deployed, how it should be deployed, and whether it is being deployed as represented.

C. IMPACT ON YOUTH

Though children possess special protections in the juvenile court system, such as different sentencing guidelines, an emphasis on rehabilitation over punishment, and criminal records that are sealed and typically expunged once they turn eighteen years of age, many continue to be criminalized by highly secretive data-driven policing technologies, tools, and programs that cause lifelong collateral consequences. These inscrutable systems have been documented to be racially skewed, are riddled with errors, and have historically included children as young as eleven years old. Moreover, users rarely notify minors of their inclusion or offer them the ability to seek their removal from such systems.

As abundant research explains, the prefrontal cortex, which controls how individuals regulate emotions, control impulsive behavior, and assess risk, does not fully develop until the age of twenty-five. Children cannot think through the long-term consequences of their behaviors, nor can they fully understand the legal process, such as *Miranda* warnings. However, many children continue to be treated as adults in the criminal legal system, in violation of their fundamental rights to special protection and to be tried by a specialized juvenile justice system.

One such system is the gang database. The proliferation of gang databases has disproportionately impacted and stigmatized Black and Latinx children. They also are disconnected from the realities of how youth live their lives. Since the inception of these databases, "police officers have been racially profiling and tracking people – primarily youth of color – suspected of 'gang involvement' often based on what they look like, where they live, and how they dress."[450]

Over three decades ago, the Los Angeles County Sheriff's Department created the first gang database in the country. Up until 2013, law enforcement agencies in California were, in fact, not required to inform minors and their parents or guardians when an officer entered a minor into a gang database.[451] These databases allow law enforcement officers to share extensive information about gangs, and to "collect, store, and analyze personal information about alleged gang members."[452] Many of them are "filled with the names and pictures of thousands of young people of color who have not been convicted of any crimes",[453] as these databases have proliferated and become more common, so have reports of the errors that they contain. With no "no clear, consistent and transparent exit process" for those on the database, it can be assumed that a vast proportion of those included were added in their teens and preteens.[454]

For example, CalGang, a database widely used in California, listed 42 infants under the age of 1 as active gang members.[455] The Chicago Police Department (CPD)'s database includes more than 7,700 people who were added to the database before they turned 18, including 52 children who were only 11 or 12 years old at the time of their inclusion.[456] An investigation published by *The Intercept* identified hundreds of children between the ages of 13 and 16 listed in the New York Police Department (NYPD)'s gang database in 2018.[457] The Boston Police Department (BPD) uses a point system to determine whether to include someone in its "Gang Assessment Database";[458] making it possible for teenagers to be designated as gang members "simply because of the people they're being seen with,"[459] and without any actual allegation of violence or criminal activity.

This data provides "disturbing insights into the police targeting of young people,"[460] and the ease with which officers can add a minor to a database for having a tattoo symbolizing a gang, for wearing clothing associated with a gang, or for repeatedly visiting "a gang area."[461] For example, in an interview with the *Gothamist,* activist Shanduke McPhatter said that he became a target for the NYPD after being added to its gang database, with no opportunity to appeal his designation as a gang member:

> With gang sweeps, you're talking about 16-year-olds, 18-year-olds, 20-year-olds, but because you put them in the gang database, and swoop them in a group, now you're taking massive amounts of people out of communities, out of families.[462]

Because of the secrecy surrounding gang databases, some have even referred to them as hidden "surveillance tool[s] for monitoring children."[463] According to Rachel Levinson-Waldman at the Brennan Center for Justice, this monitoring often takes place on social media, enabling officers to search a user's publicly available account and posts; set up an undercover account to interact with a targeted user; or use a search warrant to get additional information about a specific user.[464]

For instance, in 2014, the NYPD and the Manhattan District Attorney's Office prosecuted the largest gang case in the city's history known as the "Bronx 120" raid. The Bronx 120 raid, and the NYPD's broader anti-gang tactics, followed a 2013 class-action lawsuit that challenged the department's use of stop-and-frisk,[465] threatening to foreclose the NYPD's ability to monitor youth of color based on appearance and geography. In response, the NYPD and the Manhattan District Attorney's Office heightened their "efforts to understand, oversee, and infiltrate the digital lives of teenagers," [466] tracking the activity of children as young as 10 through social media services like Twitter, Facebook, YouTube, and Instagram. In fact, though the youngest individual charged in the Bronx 120 raid was 18 years of age, "because the conspiracy allegedly went back to 2007," the average age of 110 of the defendants was only 14 at the time that prosecutors claimed a conspiracy was formed.[467]

In addition to creating fake accounts and sparking "online friendships to sidestep privacy settings,"[468] these "social media busts" join an escalating pattern of mass arrests by the NYPD where social media evidence and conspiracy statues are used to arrest large numbers of defendants.[469] Given their brain development and their related impulsivity, children, adolescents, and young adults often do not understand the various ways their social media posts can be interpreted, and how those interpretations can lead to law enforcement surveillance, arrest, and prosecution. Accordingly, the "mining of social media posts…is a recipe for abuse." [470] In fact, as Josmar Trujillo and Alex Vitale write, "[a] social media search on any given day can find hundreds, if not thousands of posts referencing a gang, most of which are clearly not related to organized crime."[471]

This mix of adolescence, law enforcement's misinterpretations of their words and posts, and a racialized criminal legal system have proved to be toxic for many Black and Latinx youth:

> While there is a growing amount of research dedicated to deciphering how social media relates to gang violence, little, if any, has sought to separate public expressions of "gangs" to actual violence. Amongst youth, words are fluid and meant to be accessible to many. Police can, however, wittingly or unwittingly take dangerous liberties by ascribing criminality or violence to these expressions. One recent report found that police "massively overestimated the direct linkage between what someone does online and what someone does offline.[472]

Furthermore, at least one expert has questioned whether violence could have been prevented if the NYPD and the District Attorney had chosen to work with members of the community, rather than "secretly recording, watching, and amassing information."[473] As K. Babe Howell, an associate professor at CUNY School of Law, wrote:

> First, one may question the wisdom of watching, listening, spying, waiting and then using conspiracy charges to link dozens of young people to offenses committed by others instead of intervening to defuse the rivalry. Second, one may wonder how a military-style raid to accomplish regular law enforcement goals affects police-community relations. Having obtained the indictment and surveilled the individuals for years, why enter their homes wearing bulletproof vests, with firearms drawn, pointing weapons at family members, while helicopters whir overhead?[474]

Likewise, in Chicago, young men of color were overrepresented on the Strategic Subject List, which ranked individuals with a criminal record from 1 to 500, based on their probability of being involved in a shooting or murder, either as a victim or an offender,[475] on the list, as they are in the CPD's gang database. Unlike criminal registries, individuals can be listed on Chicago's gang database simply by suspicion or association, rather than only after they've committed a gang-related crime.[476] An investigation published by *ProPublica* found that the CPD's gang database was not only "riddled with errors and illogical information that police officials couldn't explain," but also that officers themselves previously raised concerns over its fairness, accuracy, and constitutionality.[477] The CPD has repeatedly declined "to give the total number of juveniles it considers to be gang members" in its system; and in her 2018 meeting with the Task Force, attorney Chaclyn Hunt suggested that it was unlikely that the CPD would ever take children off of the gang database:

One of the main reasons is that these are not records that are subject to expungement. They are investigatory tools. It's like if they had investigated you on a specific case, but you weren't listed publicly as a subject: you can't get that expunged. That doesn't disappear from the police department's files, and that information, like whether you are in a gang or on the strategic subject list -- your criminal history, all of that is available to the police anyway.

> One of the main reasons is that these are not records that are subject to expungement. They are investigatory tools. It's like if they had investigated you on a specific case, but you weren't listed publicly as a subject: you can't get that expunged. That doesn't disappear from the police department's files, and that information, like whether you are in a gang or on the strategic subject list -- your criminal history, all of that is available to the police anyway.[478]

Critics have argued that gang databases, and the methods used to obtain intelligence and data for such systems, function like "black boxes," with little information available on how someone gets on or off these lists, making them a prime tool for racial profiling.[479] In an interview with *The Intercept*, attorney Anthony Posada noted that initiatives like Operation Crew Cut were arguably intended to add Black and Latinx youth to the NYPD's expansive gang database:

> [The data] confirms to us what we were suspicious about and wanting to know more about. It is confirmation that there are these active programs ... that really destroy the ability to build community trust — that are secretive, that are unconstitutional, that label people without an ability for them to be removed from that data.[480]

Studies have also shown that once an individual is listed in a gang database, they will likely encounter increased police attention and harassment. Since gang databases make gang identification information significantly more accessible to law enforcement officers, this has resulted in the more widespread stopping of young people of color, even without suspicion of criminal activity.[481] Allegations of gang involvement have also been found to deter students in Chicago from accessing their neighborhood schools, with researchers at the University of Illinois at Chicago claiming that Chicago Public Schools can refuse to admit young people with alleged gang designations.[482]

D. IMPACT ON CONSTITUTIONAL RIGHTS AND THE CRIMINAL PROCESS

Data-driven policing has proliferated so quickly that solutions lag for the myriad constitutional rights that are implicated by its deployment and use. The aggregation and classification of vast and disparate types of personal information raises serious concerns about the First, Fourth, Fifth, Sixth, and Fourteenth Amendment rights of those suspected and accused in criminal cases. This section highlights some of the constitutional concerns that arise with the use of these data-driven technologies, starting with Justice Brandeis' dissent in *Olmstead v. United States,* where he wrote:

> The makers of our Constitution undertook to secure conditions favorable to the pursuit of happiness. They recognized the significance of man's spiritual nature, of his feelings and of his intellect. They

> knew that only a part of the pain, pleasure and satisfactions of life are to be found in material things. They sought to protect Americans in their beliefs, their thoughts, their emotions and their sensations. They conferred, as against the Government, the right to be let alone — the most comprehensive of rights and the right most valued by civilized men. To protect that right, every unjustifiable intrusion by the Government upon the privacy of the individual, whatever the means employed, must be deemed a violation of the Fourth Amendment. And the use, as evidence in a criminal proceeding, of facts ascertained by such intrusion must be deemed a violation of the Fifth. [483]

Data-driven policing creates fertile ground for the types of intrusions and violations Brandeis identified nearly a century ago, but solutions are not readily available.

1. Fourth Amendment

Data-driven policing raises serious questions for a Fourth Amendment analysis. Prior to initiating an investigative stop, law enforcement typically must have either reasonable suspicion or probable cause. Reasonable suspicion, often described as more than a hunch but less than probable cause,[484] requires that a law enforcement officer have reason to believe that a particular person is involved in criminal activity.[485] The question then becomes: to what extent should an algorithm be allowed to support a finding of probable cause or reasonable suspicion?

Does a person loitering on a corner in an identified "hotspot" translate to reasonable suspicion? What if an algorithm identified that person as a gang member or someone likely to be involved in drug dealing or gun violence? Can an algorithm alone ever satisfy the reasonable suspicion requirement? As Elizabeth Joh wrote,

> Police that rely on big data tools to identify those people and places that deserve attention are using these programs to help develop their own assessments about suspicion. These assessments in turn can help develop the legal suspicion necessary to conduct stops, frisks, and arrests. At some point in the near future, courts will have to determine whether an algorithm's determination can form the basis, at least in part, of Fourth Amendment suspicion. If informants and tips can help develop reasonable suspicion, it is likely that courts will accept big data analysis as another source of information for the police as well. The problem for courts and defendants hoping to find out how a big data program has arrived at its conclusions is that the suspicion itself has been outsourced, at least in part. How an algorithm recommended police attention to one person or city block rather than another may be guarded as a "trade secret" that the algorithm's creators are unwilling to reveal. [486]

Given the possibility that algorithmic decision-making systems (ADS) are supported by outdated, inaccurate and biased data and the inherent likelihood of false positives, Courts should not allow ADS outputs to substitute for other Fourth Amendment analysis. Moreover, if algorithms rely on information obtained from private data brokers that traditionally would require a warrant to access, then law enforcement essentially performs an end-run around the Fourth Amendment by using those algorithms to draw conclusions about an individual's suspected criminality.[487]

Unfettered access to various data sets allows police officers to incorporate a far more detailed personal profile into their reasonable suspicion database than would otherwise be available to them. Joh further explains that:

> Automating the suspicion analysis — in whole or in part — could dramatically change policing. Some information that previously would not have been known to individual officers, either because it was unknown or because it would have been too cumbersome to retrieve quickly, becomes part of the investigations process. Big data might also bring new and unexpected insights about criminal behavior. The scale of automation also widens the scope of surveillance over many more potentially suspicious persons.[488]

The result is a self-fulfilling prophecy. The mere fact of conducting stops can cause algorithms to double down on a particular area as a hotspot, and then interpret the data about the stop as a further indication of a person's dangerousness, resulting in more policing of the same neighborhoods and increased police encounters

for the same population?[489] The lack of transparency and clarity on the role that predictive algorithms play in supporting reasonable suspicion determinations could make it nearly impossible to identify a potential Fourth Amendment violation.

Concerningly, even if a defendant were able to find inaccuracies or other faults stemming from database usage in their own case, the exclusionary rule may not apply. In *Herring v. United States*, the Supreme Court held that a defendant could not suppress evidence obtained in "objectively reasonable reliance" on a recordkeeping error in a police database.[490] An officer's reliance might not be "objectively reasonable," the Court suggested, where there were "systemic errors" in the database.[491] To prove systemic errors in a data-driven policing database, of course, a defendant must first be given access to the underlying algorithms and data. Thus, *Herring* and the "trade secrets" privilege have the potential to work in tandem to divest defendants of their Fourth Amendment rights.

2. Fifth Amendment

As was detailed previously, the use of data-driven policing technologies is often not publicly disclosed. Even if the use of the technology is a matter of public record, the inputs used, training data and algorithms, are proprietary and therefore shielded from scrutiny. This raises a number of due process issues that implicate a person's right to a fair trial.

Recall Kevin Vogeltanz discovered New Orleans' use of Palantir's Gotham program almost by accident. This lack of notice and access leaves defense lawyers with no opportunity to inquire into what was behind law enforcement's actions and what led them to suspect and arrest the accused. This in turn raises serious due process concerns.

In *Brady v. Maryland*,[492] the Supreme Court found that when the government withholds evidence that is material to a determination of either guilt or punishment, it violates the due process rights of the accused. In a jurisdiction with a tool akin to Chicago's Strategic Subject List or a gang database, lack of access prevents defense lawyers from requesting information that would shed light on why their client appeared on the list and whether there were others with higher "scores" or more suspect classifications, implicating the due process rights the Court established in *Brady*. Without disclosure of the technologies involved, defense attorneys do not have the opportunity to challenge whether the results or the tools themselves were inaccurate or improperly deployed. As Michelle Fields noted to the Task Force:

> In regards to due process, as of now, no one currently on the gang database list has a right to oppose it, has a right to review it, has a right to even have an open hearing as to how they are able to now remove their names from that database. What is happening when you use that information, especially for Black and Brown communities, the collateral consequences are…barriers to housing, employment.[493]

3. Sixth Amendment

Algorithmic tools often use claims of proprietary software and trade secrets to shield their technology from outside scrutiny. The companies that develop the tools conduct their own validation studies, rather than rely on independent verification and validation, which is the accepted practice. Allowing companies with a financial interest in the success of their tools to validate their own technologies with no outside scrutiny is scientifically suspect.[494] It also frustrates any defense effort to challenge the reliability of the science underlying the novel software.

Two recent cases show that the tide is beginning to turn. In both *New Jersey v. Pickett*[495] and *United States v. Ellis*,[496] the defense requested access to a company's (TrueAllele) software source code. TrueAllele is used to conduct a probabilistic genotyping analysis for mixed DNA samples. Both courts concluded that the defense should be given access subject to a protective order. As the court in *Pickett* concluded, "anything less than full access contravenes fundamental principles of fairness, which indubitably compromises a defendant's right to present a complete defense."[497] While these tools are distinct from data-driven policing technologies, decisions establishing that software source code can be accessed in criminal cases sets a promising precedent for other technologies claiming trade secret protections.

Courts have the tools to ensure that corporate interests do not undermine constitutional rights in criminal cases. Creating these avenues for transparency and accountability will be a critical check on data-driven policy technology.

4. Fourteenth Amendment

As the Task Force heard throughout their investigation, data-driven policing tools often reinforce or even exacerbate the racial biases that have always existed in policing. The application of artificial intelligence and machine learning to questions of who is surveilled, stopped, questioned, arrested, and otherwise criminalized puts a finding of discriminatory application of the law even further out of reach. According to legal scholar Aziz Huq:

The concerns of constitutional law simply do not map onto the ways in which race impinges on algorithmic criminal justice. The result is a gap between their legal criteria and their objects.... The replacement of unstructured discretion with algorithmic precision, therefore, thoroughly destabilizes how equal protection doctrine works on the ground.

> The concerns of constitutional law simply do not map onto the ways in which race impinges on algorithmic criminal justice. The result is a gap between their legal criteria and their objects....The replacement of unstructured discretion with algorithmic precision, therefore, thoroughly destabilizes how equal protection doctrine works on the ground.[498]

This concern is particularly acute in a Fourteenth Amendment analysis where the assessment relies on discriminatory intent [499] as a threshold finding. The data-driven tools are structured in a manner by which bias is buried beneath the technology. As Vincent Southerland told the Task Force:

> If the police are shaping the data, then that data is going to shape what the police end up engaging in or what they end up doing. [...] In those instances [where] they were using the PredPol tool, ... what you have is essentially a police officer being told, 'Look, a crime is going to happen in this particular community. Be on the lookout for crime.' Almost priming them to engage in racial profiling and engage in this heightened level of suspicion of individuals who are walking around. [...] You're almost priming police to engage in this misconduct based on their own interactions with the community.[500]

Because any bias is filtered through an algorithm, critics have accused data-driven tools of "techwashing" [501] the biases inherent in the data. While machine learning is not advanced enough to formulate intent, "(i)n the policing context, the unthinking use of algorithmic instruments will reinforce historical race-based patterns of policing."[502] In order to address allegations of systemic bias in these algorithmic tools, attorneys will need to litigate the intent standard out of the Fourteenth Amendment analysis and insist on a disparate outcomes test when technology is involved.

5. First Amendment

Although First Amendment concerns are not primary in criminal prosecutions, there are several First Amendment issues raised by data-driven policing programs and technologies. When people are criminalized based on their associations and their participation on social media, they are subject to what Elizabeth Joh calls the "surveillance tax." As Joh writes, the intrusiveness of surveillance extends beyond arrest: "Knowledge of surveillance alone can inhibit our ability to engage in free expression, movement, and unconventional behavior."[503]

In *Stanford v. Texas*,[504] the Court found that Fourth Amendment protections are particularly sensitive when First Amendment rights are also implicated, as police should not be the arbiters of First Amendment protections. "The constitutional impossibility of leaving the protection of those freedoms to the whim of the officers charged with executing the warrant is dramatically underscored by what the officers saw fit to seize under the warrant in this case."

Gang designations and inclusion on lists of potential offenders are often based on proximity, associations, and social media interactions, rather than facts and evidence. The low bar for inclusion in such data bases,

the lack of notice, and the inability to challenge one's inclusion create circumstances where young people are forced to live with the potentially life-changing consequences of such designations. Jarrell Daniels posed this to the Task Force:

> At what point is a person nonaffiliated anymore? And at what point are they removed from the database system that NYPD is using? They never have answers for that. It's like, you know, I'm going to be there for the rest of my life, until I die.[505]

Even if the initial designation was based in fact and evidence, the opaque nature of inclusions in these databases and the inability to challenge that inclusion do not reflect the reality that people change over time and age out of crime. Taylonn Murphy put it this way:

> How do you know, once you are on the database, how do you get off the database? We don't even know what's going on with the database. I know I'm not the same person I was at 19 or 20 that I am now at almost 50. And I know many of you guys aren't the same person you was when you were in college or when you were in high school.[506]

E. POLICE DEPARTMENTS AND JURISDICTIONS THAT HAVE ENDED OR DECIDED NOT TO PURSUE DATA-DRIVEN POLICING

As law enforcement's use of "advanced surveillance technologies like artificial intelligence and machine learning" has increased dramatically over the past several years,[507] "so has criticism based on the lack of transparency and the potential for bias and abuse."[508] This is particularly true for technologies that rely on algorithms, such as data-driven policing, as Malkia Cyril noted in their meeting with the Task Force:

> So, the question becomes not, "Should technology have nothing to do with policing?" It's more like, "How can technology actually produce outcomes that are about justice, equity, and fairness?" And that means that, one, we'd have to turn the tables on who the technology is supposed to help. If it's there to help police officers, then it's actually going to replicate the dynamics of power and dynamics of inequity. If it's there to help those being policed, well, that might be different. And we ain't seen that yet.[509]

More and more cities have recently begun to reassess their contracts and policies for data-driven policing technologies, after facing considerable criticism from communities, activists, technologists, academics, and even the mathematicians who helped create the algorithms behind predictive policing.[510] In 2019, the *Los Angeles Times* reported that numerous police departments across the country were ending their contracts with PredPol because they determined it did not help reduce crime and "that it provided information already being gathered by officers patrolling the streets."[511] Earlier this year, a group of candidates running in the Manhattan District Attorney primary promised to "sever ties" with Palantir and other companies that "offer invasive surveillance."[512] And as of January 2021, at least 25 municipal government entities have passed legislation governing law enforcement's use of new technology.[513]

This follows a growing shift to introduce legislation at the state and local level that would prohibit police departments or other agencies from acquiring or using surveillance technologies without public input and the approval of elected representatives. For example, the ACLU proposed sample legislation titled, "Community Control Over Police Surveillance" (CCOPS), which would provide "local city councils control over the purchase and use" of such technologies, in addition to insight into the "contracts between local police departments and the private companies who develop these technologies."[514] Similar regulations have been adopted across the country, with CCOPS serving as a model in many local municipalities.[515]

CONCLUSION

The United States is at a critical and potentially galvanizing moment, with broad-based constituencies coalescing around the need to break away from racialized policing that has confined communities of color. The relationships between law enforcement and communities of color have long been broken. Efforts to "fix" these relationships and to transform police were well underway in several communities throughout the United States when Minneapolis police officers killed George Floyd on May 25, 2020, and Louisville police officers killed Breonna Taylor in the early morning hours on March 13, 2020. Those two tragedies – which follow the painful history of law enforcement killings of Black men, women, and children throughout U.S. history – have broadened the clarion calls to radically transform, defund, and dismantle law enforcement.

While transformation, defunding and dismantlement carry different meanings, the bottom-line is that the present moment presents a unique opportunity to join the broad-based efforts, anchored in impacted communities and pushed by young leaders to change history: specifically, to transform the ways in which law enforcement sees and interacts with communities of color.

Transformation in this context requires that all aspects of policing be examined, analyzed, and scrutinized. This Report does so in the context of data-driven policing technology. As Aziz Huq writes, "Adoption of machine learning within the criminal legal system changes the scale, reach, and operation of state power."[516] Not only does it expand the scope of policing, it also exacerbates racial biases that are already baked into the system, as Cathy O'Neil told the Task Force:

> Historical data, and particularly the racist history of the broken windows theory of policing, that sort of data artifact of that uneven racist policing system meant that hotspot policing, predictive policing, was just going to recreate that system. Even if we don't say we do broken windows policing anymore, we actually still do it essentially because of the data artifact that we carry with us.[517]

Given the urgent needs to protect Black and Latinx communities from racialized policing, to protect the constitutional rights of individuals during police encounters, to protect the constitutional rights of individuals charged with crimes, and to reduce mass incarceration and mass criminalization, the role of these technologies must be scrutinized.

Simply put, law enforcement's use of policing technology must help and not hurt impacted communities. Any technology used by law enforcement must uplift the community, and respect the dignity and rights of community members. Also, to ensure both fairness to individuals accused of crime and respect for their constitutional rights, any technology that police officers have used, in any way, against a person accused of a crime must be made known to the defense attorney representing the accused and the court overseeing the prosecution. The stakes are too high, and the relationships are too fragile to implement data-driving policing technologies that fall short of these principles.

In light of these principles, the evidence presented to the Task Force, and the vast and continuously emerging literature on data-driven policing that the Task Force has reviewed, the Task Force's overarching

recommendation is that the technologies and tools that support data-driven policing should not be used. Police departments that are currently using these technologies should cease. For the most part, both private companies and police departments have already developed and implemented many of these tools in secrecy, without consulting with or even informing impacted communities, and with little to no oversight. This was further emphasized by Malkia Cyril in their meeting with the Task Force:

> The thing about predictive policing technologies is, again, people don't know they're being used. You don't know when they're being used. You don't actually even know how they influence the decisions that police officers are making....It doesn't matter what the specific tool is that's being deployed. The fact is that the dynamics of the over-policing continue, period; and any tool that's introduced into that dynamic will further it.[518]

Police departments have prioritized imposing these tools over forging the collaborative relationships with communities that are necessary for policing reform and for effective policing. As Taylonn Murphy expressed to the Task Force, police departments need to prioritize "community building,"[519] rather than, as Scott Levy explained, "flooding" communities with this type of policing just to further "[destabilize] the bonds and trust that you need to actually build community power."[520]

These tools broaden the net that hyper-criminalizes poor communities of color, especially Black communities, in ways that further erode the dignity of individuals and families within these communities. They also epitomize what Dr. Ruha Benjamin has referred to as "The New Jim Code" in which the "employment of new technologies...reflect and reproduce existing inequities but...are promoted and perceived as more objective or progressive than the discriminatory systems of a previous era."[521] In addition, the criminal legal system is woefully unprepared to address these tools in the context of the due process and other constitutional rights owed to individuals charged with crimes, or is otherwise uninterested in doing so. Overall, these tools suffer from a lack of transparency, a lack of analysis, and a lack of accountability. They should not be used.

However, the Task Force is mindful of the blunt reality that big data tools and technologies are proliferating in police departments throughout the country. Technology and algorithms used to predict, infer, guess, assess, and short-circuit are emerging rapidly in almost every stage of the criminal legal system, from arrest, to bail, and to sentencing.[522] While the harms and ineffectiveness of data-driven policing should cause police departments to take deep pause and deploy other methods and strategies to serve and protect communities, the Task Force anticipates that several police departments, other government entities, private companies, and various decision-makers would assert, in essence, the need to continue to speed in the other direction. Accordingly, the Task Force sets forth several alternative recommendations.

To arrive at these recommendations, the Task Force focused on the stated goals of the various technologies being used, evidence of their effectiveness, and their potential shortcomings. The Task Force's recommendations ultimately revolved around the implications of predictive policing algorithms and data-driven policing technologies on how individuals are surveilled, investigated, charged, and prosecuted, specifically addressing the impact of racial profiling, policing, and prosecution on historically overpoliced groups and communities of color.

TASK FORCE RECOMMENDATIONS ON DATA-DRIVEN POLICING TECHNOLOGIES

1. Top-Line Recommendation

Police departments must not utilize data-driven policing technologies[523] because they are ineffective; lack scientific validity; create, replicate and exacerbate "self-perpetuating cycles of bias"[524]; deeply entrench existing inequities in the system; hyper-criminalize individuals, families, and communities of color; and divert resources and funds from communities that should be allocated towards social services and community-led public safety initiatives.

While the Task Force believes these technologies should never be used, it is clear that these technologies are being considered or have been implemented in cities and towns across the country. Lack of access and transparency will hamper defense lawyers' ability to properly represent their clients. The following recommendations are for areas that are already using these technologies. These recommendations are in no way intended to serve as principles for implementing such technologies. Rather, they are mitigation efforts intended to ensure the most transparency and equity for people ensnared by these technologies, and to give defense attorneys the notice and transparency they need to defend their clients.

2. Governing Use

Police departments seeking these tools must not adopt any data-driven policing technology without first meaningfully engaging the communities where it would be deployed and without first securing approval for the technology from the elected governing bodies that represent the impacted communities.

This process must include the residents of the communities where the data-driven policing technologies would be deployed, community organizations, organizations focused on youth from the impacted communities, and attorneys with expertise in upholding the constitutional rights and civil liberties of residents from impacted communities.

As part of engaging impacted communities about the proposed data-driven technology, resources must be allocated to local governing bodies to host forums to present and describe the proposed law enforcement technology to the residents of the impacted communities. These forums would also provide a space for impacted communities and law enforcement to discuss the law enforcement need for the proposed technology, detailing how the policies governing the use, scope, and limitations of the technologies would be implemented within the defined law enforcement need. Resources and space should also be allocated to enable and empower community members to provide feedback about the technology, and to address community concerns about transparency, racial bias, and the impact of the proposed technology on civil liberties and constitutional rights. If there is a majority consensus by state or local governments and impacted communities that the proposed technology should not be used by law enforcement, then the technology should be prohibited.

3. Transparency

Prior to implementing any data-driven policing technology, law enforcement must adopt written policies governing the technology's use. Before adopting these policies, law enforcement departments must make draft

policies available to the public, provide the public with opportunities to comment on the draft policies orally and/or in writing, and incorporate public comments into the final policies. For any technology already in use but lacking such policies, law enforcement departments should immediately implement clear public policies that detail the parameters, requirements, and conditions of use.

Tech companies and developers of data-driven policing technologies have asserted trade secret evidentiary privileges as reason to deny defense discovery requests and subpoenas.[525] To facilitate transparency and avoid the exclusion of highly probative evidence,[526] companies that create and supply data-driven policing technology must waive, or otherwise not assert, claims of "trade secret privilege."[527] They must also disclose the methodologies used to build the technology to law enforcement, the impacted communities where law enforcement departments intend to deploy the technology, the legislative bodies that represent the impacted communities, and the attorneys within the jurisdiction who specialize in criminal defense and civil liberties to ensure that the technologies are scientifically sound, are employed as intended, and are limited in scope to meet the articulated law enforcement need.

Any data-driven policing technologies that are used should undergo validation studies that allow them to be subjected to a *Daubert* or *Frye* analysis. As matters of constitutional due process rights guaranteed by the Fifth and Fourteenth Amendments, all individuals must be notified of their presence on data-driven databases that law enforcement departments access and utilize, including gang databases, strategic subject lists, and other data collected through social media monitoring. These individuals must also be provided the opportunity, through a private attorney or, if they cannot afford an attorney, an appointed attorney, to challenge their inclusion on such databases, the data accumulated from the databases, and law enforcement's interpretation of the data, as well as to seek removal from the databases.

All individuals, in accordance with constitutional due process rights guaranteed by the Fifth and Fourteenth Amendments, must be notified of their removal from any data-driven databases that law enforcement departments access and utilize, including, but not limited to, gang databases and strategic subject lists.

4. Race Equity

The analysis that jurisdictions undertake when considering whether to adopt any data-driven policing technology must be conducted through a race equity lens and include a racial impact statement. The "racial equity impact assessment"[528] must be conducted by experts trained in institutional and structural racism, as well as the history of racialized policing. These experts should work with legislators, law enforcement, and community members to examine the racialized impact of the proposed data-driven policing technology. If the racial equity impact assessment of the proposed data-driven policing technology concludes that use of the technology would harm the impacted community, the technology should be prohibited from use by law enforcement.

5. Accountability

If, through the processes detailed in Recommendations #2, #3, and #4, data-driven policing technologies have been approved, law enforcement departments must adopt and issue written protocols ensuring integrity and accountability, to ensure that the departments and the impacted communities can continuously monitor and otherwise gauge the use and effectiveness of these technologies.

Integrity and accountability measures must include data-keeping, annual departmental reports on the use and accuracy of the technology, measuring and evaluating the effectiveness of the technologies through auditing, and, based on the results of these accountability measures, determining whether the use of the technology should be modified or discontinued. All reports, evaluations, data, and accountability measures produced in relation to data-driven policing technologies should be made available to the public.

6. Resources and Access for Defense Attorneys

In accordance with the constitutional rights to discovery and confrontation guaranteed by the Sixth Amendment, prosecutors must provide to defense counsel notice and a description of data-driven policing

technology that law enforcement has employed or has otherwise relied upon in the case, as well as any data based on the technology that the officers relied upon, assessed, or otherwise used in relation to the accused, including *Brady* material and any other data accumulated against the accused. Defense counsel must then be afforded time and resources to engage experts to analyze and interpret the data.

Defense lawyers must receive notice and training regarding the data-driven policing technologies employed by law enforcement departments in their jurisdictions, including the federal and state constitutional rights implicated by the technologies.

Defense lawyers should collaborate with other attorneys, technologists, and experts who understand the data-driven policing technologies employed against their clients, and should seek to incorporate law enforcement's use of the relevant tool(s) against their clients into all aspects of their representation.

Defense lawyers must have access to data-driven technology experts who can break down the technologies and consult on defense strategies vis-à-vis the data-driven tools that law enforcement relied upon to suspect, surveil, approach, or arrest, or otherwise employed against the accused.

Resources for public defenders and court-appointed counsel must be increased to respond to data-driven policing technologies in order to meet their constitutional obligation to provide zealous representation to clients impacted by these technologies.

7. Courts

Courts and prosecutors must be trained annually on the data-driven policing technologies employed by law enforcement departments, including the federal and state constitutional rights implicated by the technologies.

Judges must assess the reliability of a data-driven policing technology employed against the accused before determining whether it justified a Fourth Amendment intrusion. Data-driven technology must not form part of an officer's calculation of reasonable suspicion, unless the technology can be shown through typical evidentiary burdens that it is reliable.

Law enforcement authorities cannot utilize or otherwise rely upon data-driven technologies, such as gang databases, in any way that infringes upon the right to association guaranteed by the First Amendment.

8. Children and Youth

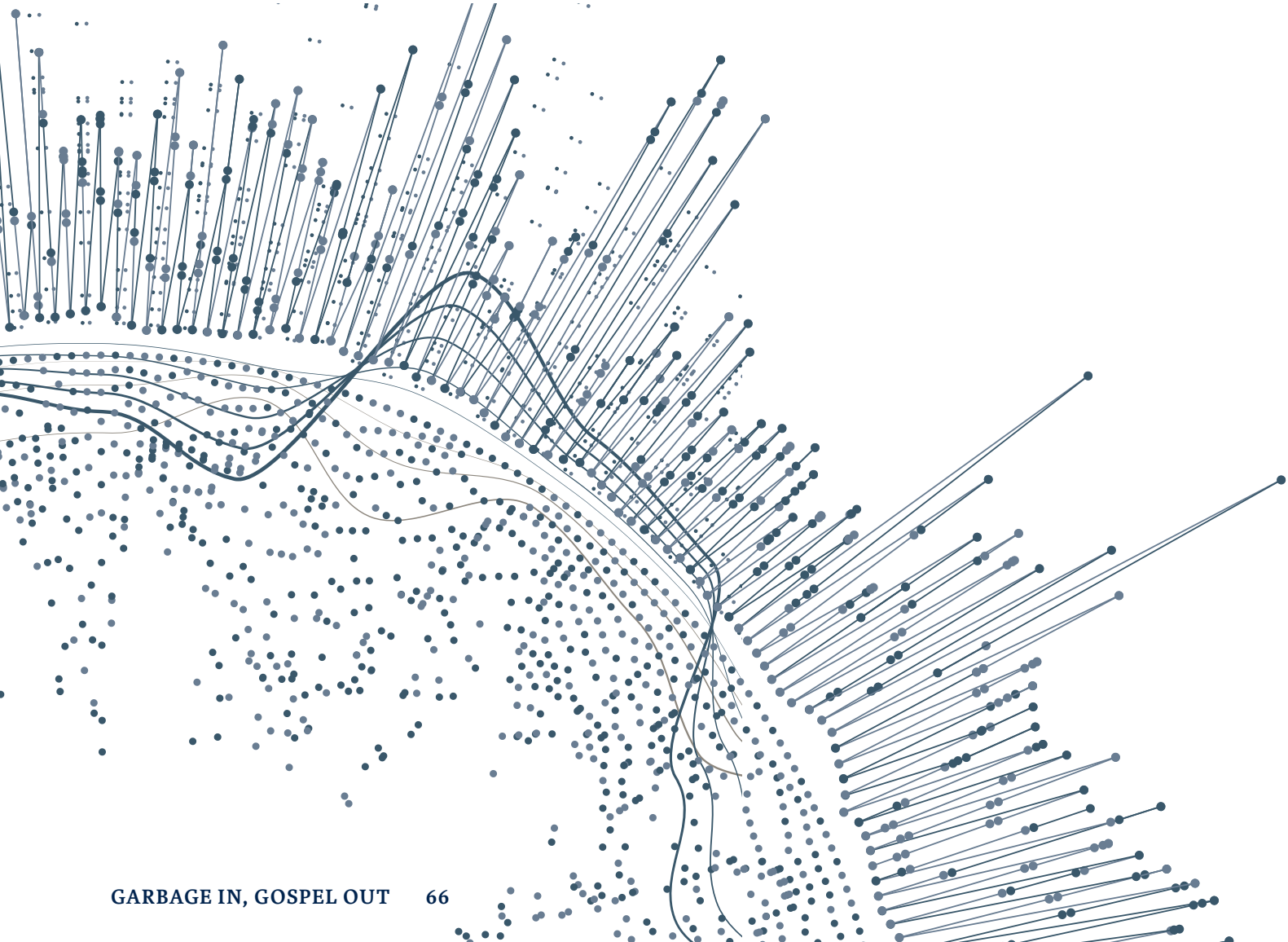
State and local jurisdictions must enact laws, policies, and protocols that protect the federal constitutional rights, state constitutional rights, and dignity interests of children and youth who are implicated or otherwise at risk of being criminalized by data-driven policing technology.

Law enforcement authorities should not include children under the age of 18 on any law enforcement database, or otherwise accumulate or access data specific to children under the age of 18 through social media monitoring or other data gathering practices.

Young people between the ages of 18 and 25 are especially vulnerable, disproportionately included on data-driven policing databases,[529] and therefore must be provided notice of their presence on any databases that law enforcement departments access and utilize, including gang databases, strategic subject lists, and other databases that incorporate social media monitoring. Individuals must be provided the opportunity, through a private attorney or, if they cannot afford an attorney, an appointed attorney, to challenge their inclusion on such databases, the data accumulated, and law enforcement's interpretation of the data, and, also, to seek removal from the databases.

An individual's ability to challenge their designation and inclusion on such databases, the data accumulated, and law enforcement's interpretation of the data should be ongoing, particularly given the impact of law enforcement interactions with children and youth on their personal development, self-esteem, and educational outcomes — including school attendance, suspensions, expulsions, and matriculation — and the correlation between these factors and involvement with the juvenile and criminal legal systems.

Any data, records, or other information contained in any law enforcement database through any data-driven policing technology and/or social media monitoring should be sealed and purged when the individual reaches 25 years of age, at which point the adolescent brain is fully formed.[530]



APPENDIX

A. OVERVIEW OF TASK FORCE MEETINGS AND WITNESSES[531]

Task Force Call with Professor Andrew Ferguson (December 8, 2017)

- Andrew Ferguson, Professor of Law, University of the District of Columbia.

Task Force Meeting in Washington, D.C. (December 18, 2017)

- Sarah Brayne, Assistant Professor of Sociology, University of Texas at Austin.
- Andrew Ferguson, Professor of Law, University of the District of Columbia.
- Daniel Kahn Gillmor, Senior Staff Technologist, ACLU Speech, Privacy, and Technology Project.
- Rachel Levinson-Waldman, Deputy Director, Liberty & National Security Program, Brennan Center for Justice at NYU School of Law. *(Previously: Senior Counsel, Liberty & National Security Program, Brennan Center for Justice at NYU School of Law.)*
- David Robinson, Visiting Scientist, AI Policy and Practice Initiative, Cornell University's College of Computing and Information Science. *(Previously: Managing Director, Upturn.)*
- Jay Stanley, Senior Policy Analyst, ACLU Speech, Privacy, and Technology Project.

Task Force Meeting in Los Angeles, CA (February 8, 2018)

- Zach Friend, Second District Supervisor, Santa Cruz County Board of Supervisors.
- Jamie Garcia, Organizer, Stop LAPD Spying Coalition.
- Hamid Khan, Campaign Coordinator, Stop LAPD Spying Coalition.
- Brian McDonald, Chief Executive Officer, Geolitica.
- John Patzakis, Chief Legal Officer, X1. *(Previously: Executive Chairman of the Board, X1.)*
- Myla Rahman, District Manager to Senator Steven Bradford
- John Raphling, Senior Researcher, U.S. Program, Human Rights Watch.
- Jessica Saunders, Research Director, Council of State Governments Justice Center. *(Previously: Senior Policy Researcher, RAND Corporation.)*
- Pete White, Executive Director, Los Angeles Community Action Network.

Task Force Call with Kristian Lum (March 29, 2018)

- Kristian Lum, Assistant Research Professor, Department of Computer and Information Science, School of Engineering and Applied Science, University of Pennsylvania. *(Previously: Lead Statistician, Human Rights Data Analysis Group.)*

Task Force Meeting in New York, NY (April 18, 2018)

- Henrik Chulu, Freelance Journalist and Digital Security Consultant.
- Cynthia Conti-Cook, Tech Fellow, Ford Foundation. *(Previously: Staff Attorney, Special Litigation Unit, Legal Aid.)*
- Michelle Fields, Co-Supervising Attorney, Community Justice Unit, Legal Aid Society.
- Yung-Mi Lee, Supervising Attorney, Brooklyn Defender Services.
- Scott Levy, Chief Policy Counsel, Bronx Defenders. *(Previously: Special Counsel to the Criminal Defense Practice, Bronx Defenders.)*
- Chad Marlow, Senior Advocacy and Policy Counsel, ACLU. *(Previously: Advocacy and Policy Counsel, ACLU.)*
- Matt Mitchell, Tech Fellow, Ford Foundation. *(Previously: Hacker, CryptoHarlem.)*

- **Taylonn Murphy, Sr.**, Community Organizer and Activist.
- **Anthony Posada**, Co-Supervising Attorney, Community Justice Unit, Legal Aid Society.
- **Jeffrey Ratcliffe**, Professor of Criminal Justice, Temple University.
- **Rashida Richardson**, Visiting Scholar, Rutgers Law School and Rutgers Institute for Information Policy and Law. *(Previously: Director of Policy Research, AI Now Institute at NYU School of Law.)*
- **Vincent Southerland**, Assistant Professor of Clinical Law and Director of the Criminal Defense and Re-Entry Clinic at NYU School of Law.
- **Josmar Trujillo**, Writer and Organizer.
- **Stephanie (Ueberall) Shaw**, Project Manager, Council of State Governments Justice Center. *(Previously: Director of Violence Prevention, Citizens Crime Commission of New York City.)*
- **Rebecca Wexler**, Assistant Professor of Law, University of California-Berkeley, School of Law. *(Previously: Visiting Fellow, Information Society Project, Yale Law School.)*

Task Force Meeting in Chicago, IL (June 7, 2018)

- **Jeremy Heffner**, Data Scientist, CentralSquare Technologies. *(Previously: Product Manager and Senior Data Scientist for Hunchlab, Azavea.)*
- **Aziz Huq**, Professor of Law, University of Chicago Law School.
- **Chaclyn Hunt**, Civil Rights Attorney, Invisible Institute.
- **Freddy Martinez**, Policy Analyst, Open the Government. *(Previously: Executive Director, Lucy Parsons Lab.)*

Task Force Call with Cathy O'Neil (October 17, 2018)

- **Cathy O'Neil**, Author, Mathematician, Data Scientist.

Task Force Meeting in San Francisco, CA (February 11, 2019)

- **Shahid Buttar**, Attorney and 2020 Candidate for California's 12th Congressional District. *(Previously: Director of Grassroots Advocacy, Electronic Frontier Foundation.)*
- **Matt Cagle**, Technology and Civil Liberties Attorney, ACLU of Northern California.
- **Malkia Devich-Cyril**, Senior Fellow and Founding Director, MediaJustice. *(Previously: Executive Director, Center for Media Justice)*
- **Brian Hofer**, Chair, City of Oakland Privacy Advisory Commission.
- **Nitin Kohli**, PhD Candidate, School of Information, University of California-Berkeley. *(Previously: PhD Student, School of Information, University of California-Berkeley.)*
- **Steven Renderos**, Executive Director, MediaJustice. *(Previously: Campaign Director, Center for Media Justice.)*
- **Philip Stark**, Associate Dean, Division of Mathematical and Physical Sciences, University of California-Berkeley.

Task Force Meeting in Washington, D.C. (May 13-14, 2019)

- **Jarrell Daniels**, Open Society Youth Activist Fellow and Justice-in-Education Scholar, Columbia University.
- **Andrew Ferguson**, Professor of Law, University of the District of Columbia.
- **Elizabeth Joh**, Professor of Law, University of California-Davis, School of Law.
- **Kevin Vogeltanz**, Attorney, Law Office of Kevin Vogeltanz, LLC.

Task Force Call with Jeffrey Brantingham (June 27, 2019)

- **Jeffrey Brantingham**, Professor of Anthropology, University of California-Los Angeles.

Task Force Call with Sean Malinowski (August 14, 2019)

- **Sean Malinowski**, Director of Policing Innovation and Reform, University of Chicago Crime Lab. *(Previously: Deputy Chief, LAPD.)*

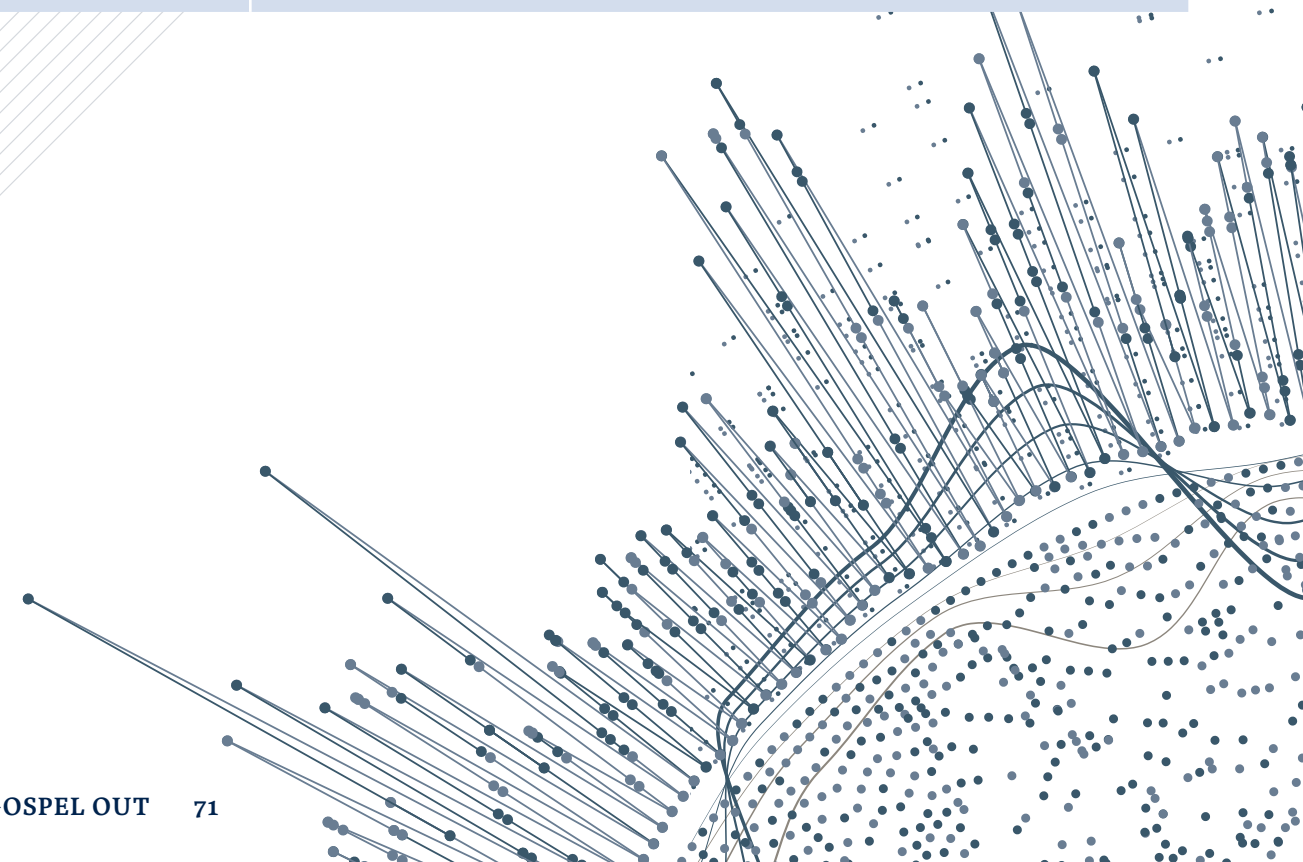
B. OVERVIEW OF STATE AND LOCAL LEGISLATION

Multiple cities and other localities have passed bills requiring transparency from law enforcement around their use of new surveillance tools or technologies. These bills, known as Community Control Over Police Surveillance (CCOPS) bills, typically impose several requirements for departments seeking to purchase or acquire new surveillance technologies.[532] First, most require that, with respect to each potential new surveillance technology, the department prepare an impact report and use policies. Second, almost all require elected bodies (e.g., city councils) approve of the purchasing or acquisition of any new surveillance technology, often based on an evaluation of the impact report and use policies. Third, many jurisdictions require annual reports on approved surveillance technologies that provides details about each technology's use. Fourth, several jurisdictions created enforcement mechanisms, including by conferring a private right of action for violations, and there are three localities that created a suppression remedy. A list of municipalities that have adopted these ordinances and relevant provisions and requirements appears below:

MUNICIPALITY/ JURISDICTION	PROVISIONS AND REQUIREMENTS
San Francisco, Cal.[533]	• Requires impact reports, use policies, and annual reports • Requires Board of Supervisors approval • Bans facial recognition technology • Confers private right of action for violations
San Francisco Bay Area Rapid Transit District, Cal.[534]	• Requires impact reports, use policies, and annual reports • Requires Board of Directors approval • Limited private right of action for violations
Oakland, Cal.[535]	• Requires impact reports, use policies, and annual reports • Requires review and recommendation by privacy commission and City Council approval • Bans facial recognition technology • Confers private right of action for violations
Berkeley, Cal.[536]	• Requires acquisition reports, use policies, and annual reports • Requires review and recommendation by Police Review Commission and City Council approval • Bans facial recognition technology
Davis, Cal.[537]	• Requires impact reports, use policies, and annual reports • Requires City Council approval • Confers private right of action for violations
Palo Alto, Cal.[538]	• Requires "surveillance evaluations," use policies, and annual reports • Requires City Council approval

San Diego, Cal.[539]	• Requires impact reports, use policies, and annual reports • Creates a Privacy Advisory Board to review and make recommendations about proposed surveillance technologies • Requires community meetings with opportunities to comment on the proposed technologies • Requires City Council approval • Confers a private right of action for violations
Santa Clara County, Cal.[540]	• Requires impact reports, use policies, and annual reports • Requires Board of Supervisors approval • Limited private right of action for violations
New Orleans, La.[541]	• Bans facial recognition technology, predictive policing technology, cell-site simulators, and characteristics tracking systems • Mandates the creation of procedures to review the use of "automated decision systems" "through the lens of equity, fairness, transparency, and accountability" • Does not otherwise require approval from City Council • Suppression remedy available for violations
Cambridge, Mass.[542]	• Requires impact reports, use policies, and annual reports • Requires City Council approval • Bans facial recognition technology • Confers private right of action for violations
Lawrence, Mass.[543]	• Requires impact reports, use policies, and annual reports • Requires City Council approval • Confers private right of action for violations • Suppression remedy available for violations
Somerville, Mass.[544]	• Requires impact reports, use policies, and annual reports • Requires City Council approval • Confers private right of action for violations • Suppression remedy available for violations
Grand Rapids, Mich.[545]	• Requires use policies and protocols • Requires City Commission approval
New York, N.Y.[546]	• Requires impact reports, use policies, and annual audits by Inspector General • Requires public comment on any proposed technology before submission to the City Council and Mayor • Does not require City Council approval

Yellow Springs, Ohio[547]	• Requires use policies and annual reports • Requires Village Council approval
Pittsburgh, Penn.[548]	• Requires impact and use policies for "predictive policing technology" and "facial recognition technology" • Requires City Council approval • Does not impact data available through other government entities or intergovernmental agreements
Nashville, Tenn.[549]	• Requires Metropolitan Council approval for new surveillance technologies to be used on any "public right-of-way" • Bans license plate scanners
Seattle, Wash.[550]	• Requires impact reports that include use policies and annual reports • Requires annual "equity impact assessment" • Requires community meetings with opportunities for public comment • Requires City Council approval • Creates a Community Surveillance Working Group that provides independent impact reports • Limited private right of action for violations
Madison, Wis.[551]	• Requires use policies and annual reports • Requires Common Council approval • Permits the mayor and Common Council to require resident feedback and comment for selected technologies



C. OVERVIEW OF POLICE DEPARTMENTS THAT HAVE SUSPENDED OR TERMINATED CONTRACTS WITH DATA–DRIVEN POLICING PROGRAMS

POLICE DEPARTMENT	TERMINATION OR SUSPENSION OF DATA–DRIVEN POLICING
Santa Cruz Police Dept.	The Santa Cruz Police Department, which began predictive policing with a pilot project in 2011, had previously placed a moratorium on the practice in 2017, and the city ordinance implemented in 2020 bans the practice permanently.[552]
Oakland Police Dept.	The Oakland City Council voted unanimously to ban the use of biometric technology and predictive policing technology in January 2021, and it is the first city in the nation to put such bans in place.[553]
New Orleans Police Dept.	The New Orleans City Council passed an ordinance in December 2020 that regulates certain parts of the city's surveillance system and places an outright ban on specific pieces of surveillance technology, including facial recognition software and predictive policing.[554]
Hagerstown Police Dept.	The Hagerstown Police Department canceled its $15,000-a-year software service in 2018 after a study commissioned by the department found that "crimes reported at a police station had skewed the data and predictions," in addition to lacking effectiveness.[555]
Pittsburgh Police Dept.	In June 2020, the City of Pittsburgh suspended its predictive policing program due to alleged concerns about racial bias, and the current ordinance states that the public safety department cannot obtain, retain, access, or use neither facial recognition technology nor predictive policing technology.[556]
Milpitas Police Dept.	The Milpitas Police Department terminated their contract with the predictive policing program, Geolitica, one year into their three-year, $37,000 contract because "the minimal benefit did not justify continuing costs."[557]
Rio Rancho Police Dept.	The Rio Rancho Police Department terminated their contract with the predictive policing program Geolitica because it "never panned out," "it didn't make much sense to [the department]", and because "it wasn't telling anything [the department] didn't know."[558]
Mountain View Police Dept.	The Mountain View, Calif., Police Department discontinued their contract with the predictive policing program, Geolitica, after spending more than $60,000 on the program between 2013 and 2018 because the "results were mixed."[559]
Palo Alto Police Dept.	After three years, the Palo Alto Police Department suspended their contract with the predictive policing program, Geolitica, because they "didn't find it effective," "[they] didn't get any value of it," and "it didn't help [the department] solve crime."[560]

ENDNOTES

1. The mission of NACDL's Task Force on Predictive Policing was to study the issues surrounding the use of various data collection tools and analytical techniques that purport to prospectively identify where criminal activity is likely to occur and the people likely to be involved. The Task Force evaluated the impact of these techniques on privacy and other individual constitutional rights, issued recommendations for best practices to safeguard those rights, and provided legal and technical assistance to educate defense practitioners in addressing the use of these tools and techniques. To arrive at these recommendations, the Task Force focused on the stated goals of the various technologies being used, evidence of their effectiveness, and their potential shortcomings. The overarching principle driving the Task Force's evaluation and recommendations were the implications of data-driven policing on how individuals are surveilled, investigated, charged, and prosecuted. The Task Force specifically addressed the implications of data-driven policing for racial profiling and the impact of policing and prosecution on historically overpoliced groups and communities of color.

2. "Our flagship software, the Palantir Platform, features a full suite of analytical tools that enable organizations and their analysts to collaboratively generate actionable insights from large and disparate data sets. The Palantir Platform has been used to tackle data-driven problems in myriad contexts, from efficiently delivering aid to victims of a natural disaster to helping law enforcement agencies (LEAs) coordinate efforts to track a dangerous fugitive or rescue an abducted child." *See* Palantir Technologies, *Palantir and Law Enforcement: Protecting Privacy and Civil Liberties,* https://web.archive.org/web/20170530065140/http://www-01.ibm.com:80/software/data/bigdata/

3. "The NYPD's precision policing consists of two important components: an intelligence-led investigative component and a neighborhood coordination component. The intelligence-led investigative side of this policy uses predictive analysis of all crimes to identify the small population of criminals who commit most of the violent crimes. [1] It also identifies small areas on the map with a higher intensity of crime, otherwise known as hot spots. [2] This new approach targets illegal firearms, precisely identifies gangs, and establishes a database for recidivists. The second component connects citizens with the police through an initiative known as neighborhood coordination. This program has three core goals: to reduce crime further, to promote trust and respect, and to solve problems by collaborating with residents. Many small community concerns are being solved collectively rather than through strict enforcement of minor crimes." See Muhammad Ashraf, Precision Policing: A Way Forward to Reduce Crime (2020), https://www.hsaj.org/articles/16249.

4. Modern day policing emerged from the shadows of slave patrols and night-watch systems in the 1800s. As a profession, policing began in the mid-1800s. Boston is credited with forming the first full-time police force in the United States in 1838. Other states and cities soon followed Boston's lead and formed their own police departments. Sally Hadden, *Slave Patrols: Law and Violence in Virginia and the Carolinas* (2003).

5. *Bringing Big Data to the Enterprise*, IBM, http://www-01.ibm.com:80/software/data/bigdata/.

6. Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 Wash. L. Rev. 35, 40 (2014). [hereinafter *Policing By Numbers*].

7. *Id.*

8. Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing* (October 26, 2015). 10 Harv. L. & Policy Rev. 15 (2015) [hereinafter *The New Surveillance Discretion*].

9. *Id.*

10. Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 Am. L. Reg 327, 330 (2015) [hereinafter *Big Data and Predictive Reasonable Suspicion*].

11. Lindsey Barrett, *Reasonably Suspicious Algorithms: Predictive Policing at the United States Border,* 41 N.Y.U. Rev. of L. & Soc'y 327, 334 (2018).

12. Walter L. Perry et al., *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, RAND Corporation (2013).

13. Andrew Guthrie Ferguson, *Predictive Policing Theory*, 24 Cambridge Handbook of Policing in the U.S. 492 (ed. Tamara Rice Lave & Eric J. Miller) (2019) [hereinafter *Predictive Policing Theory*].

14. Upturn, *Stuck in a Pattern* (2016), https://www.upturn.org/reports/2016/stuck-in-a-pattern/.

15. Andrew G. Ferguson, *Policing Predictive Policing,* 94 WASH. U. L. REV. 1009, 1125 (2017) [hereinafter *Policing Predictive Policing].*

16. *Policing By Numbers supra* note 6, at 44.

17. Molly Griffard, *A Bias-Free Predictive Policing Tool?: An Evaluation of the NYPD's Patternizr*, 47 Fordham Urb. L.J. 43, 55 (2019).

18. Kristian Lum & William Isaac, *To Predict and Serve?,* 13 SIGNIFICANCE 14, 17 (2016).

19. *Id.*

20. *Id.* at 16.

21. *Id.*

22. Barrett, *supra* note 11, at 337.

23. *Id.* at 341.

24. Elizabeth E. Joh, *Feeding the Machine: Policing, Crime Data, & Algorithms*, 26 Wm. & Mary Bill Rts. J. 287, 300 (2017) [hereinafter *Feeding the Machine*].

25. Griffard, *supra* note 17, at 52.

26. Barrett, *supra* note 11, at 340.

27. *Id.* at 344.

28. Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System* (February 21, 2017). 70 Stanford Law Review 1343 (2018), 1368.

29. *Id.*

30. Task Force Interview with Professor Elizabeth Joh at 210, UC Davis School of Law, in Washington D.C. (May 5, 2019). (Dr. Joh appeared via video teleconference).

31. Dacia Anderson, *Un-handcuffing Minors from the Gang Life*, 45 McGeorge L. Rev. 575, 575 (2014).

32. Youth Justice Coal., *Campaign Research: Gang Injunctions and Gang Data Base* 6, *available at* https://www.njjn.org/uploads/digital-library/resource_263.pdf [hereinafter *Gang Injunctions and Data*] (last visited July 3, 2013).

33. Ana Muniz, *What's Wrong with California's Gang Databases and Gang Injunctions,* Open Society Foundation, (June 22, 2013), https://www.opensocietyfoundations.org/voices/whats-wrong-californias-gang-databases-and-gang-injunctions.

34. Salvador Hernandez, *A Database of Gang Members in California Included 42 Babies*, Buzzfeed News, Aug. 11, 2019, https://www.buzzfeednews.com/article/salvadorhernandez/database-of-gang-members-included-42-babies

35. Youth Justice Coal., *Tracked and Trapped – Youth of Color, Gang Databases, and Gang Injunctions* (Dec. 2012), https://www.youth4justice.org/wp-content/uploads/2012/12/TrackedandTrapped.pdf.

36. *Id.* at 5.

37. Alice Speri, *N.Y. Gang Databases Expanded by 70% Under Mayor Bill De Blasio*, The Intercept (June 11, 2018), https://theintercept.com/2018/06/11/new-york-gang-database-expanded-by-70-percent-under-mayor-bill-de-blasio/ [hereinafter *N.Y. Gang Databases Expanded by 70% Under Mayor Bill De Blasio*].

38. *ACLU Demands Records on Boston's "Gang Database" Used in Deportations*, ACLU Mass., (Nov. 15, 2018 at 9:30 AM), https://www.aclum.org/en/news/aclu-demands-records-bostons-gang-database-used-deportations.

39. Shannon Dooling, *Here's What we Know About Boston's Gang Database*, WBUR, (Jul. 26, 2019), https://www.wbur.org/news/2019/07/26/boston-police-gang-database-immigration

40. The Policing in Chicago Research Grp., *Tracked and Targeted: Early Findings on Chicago's Gang Database* (Feb. 2018), http://erasethedatabase.com/wp-content/uploads/2018/02/Tracked-Targeted-0217.pdf [hereinafter *Tracked and Targeted*].

41. Anderson, *supra* note 31, at 582.

42. *Id.* at 582.

43. Rachel Levinson-Waldman, *Private Eyes, They're Watching You: Law Enforcement's Monitoring of Social Media*, 71 Okla. L. Rev. 997 (2019).

44. Emmanuel Felton, *Gang Databases are a Life Sentence for Black and Latino Communities*, Pacific Standard (Mar. 15, 2018), https://psmag.com/social-justice/gang-databases-life-sentence-for-black-and-latino-communities.

45. *Tracked and Trapped, supra* note 35.

46. The Fourth Amendment permits brief investigative stops — such as the traffic stop in this case — when a law enforcement officer has "a particularized and objective basis for suspecting the particular person stopped of criminal activity." United States v. Cortez, 449 U.S. 411, 417– 418 (1981); see also Terry v. Ohio, 392 U. S. 1, 21–22 (1968). The "reasonable suspicion" necessary to justify such a stop "is dependent upon both the content of information possessed by police and its degree of reliability." Alabama v. White, 496 U. S. 325, 330 (1990). The standard takes into account "the totality of the circumstances — the whole picture." Cortez, *supra*, at 417. Although a mere "hunch" does not create reasonable suspicion, Terry, *supra,* at 27, the level of suspicion the standard requires is "considerably less than proof of wrongdoing by a preponderance of the evidence," and "obviously less" than is necessary for probable cause, United States v. Sokolow, 490 U. S. 1, 7 (1989). Navarette v. California, 572 U.S. 393 (2014).

47. "The substance of all the definitions" of probable cause "is a reasonable ground for belief of guilt." McCarthy v. De Armit, 99 Pa. 63, 69, quoted with approval in the *Carroll* opinion. 267 U.S. at page 161, 45 S.Ct. at page 288, 69 L.Ed. 543, 39 A.L.R. 790. And this "means less than evidence which would justify condemnation" or conviction, as Marshall, C.J., said for the Court more than a century ago in Locke v. United States, 7 Cranch 339, 348, 3 L.Ed. 364. Since Marshall's time, at any rate, it has come to mean more than bare suspicion: Probable cause exists where "the facts and circumstances within their (the officers') knowledge and of which they had reasonably trustworthy information (are) sufficient in themselves to warrant a man of reasonable caution in the belief that' an offense has been or is being committed." Carroll v. United States, 267 U.S. 132, 162, 45 S.Ct. 280, 288, 69 L.Ed. 543, 39 A.L.R. 790. See Brinegar v. United States, 338 U.S. 160.

48. "Police that rely on big data tools to identify those people and places that deserve attention are using these programs to help develop their own assessments about suspicion. These assessments in turn can help develop the legal suspicion necessary to conduct stops, frisks, and arrests. At some point in the near future, courts will have to determine whether an algorithm's determination can form the basis, at least in part, of Fourth Amendment suspicion. If informants and tips can help develop reasonable suspicion, it is likely that courts will accept big data analysis as another source of information for the police as well. The problem for courts and defendants hoping to find out how a big data program has arrived at its conclusions is that the suspicion

itself has been outsourced, at least in part. How an algorithm recommended police attention to one person or city block rather than another may be guarded as a 'trade secret' that the algorithm's creators are unwilling to reveal." *See* Elizabeth Joh, *The Undue Influence of Surveillance Technology Companies on Policing,* 91 N.Y.U. L. Rev. Online 101, 134 (2017) [hereinafter *The Undue Influence of Surveillance Technology Companies on Policing]*

49. 373 U.S. 83 (1963)

50. Unvalidated and unreliable forensic evidence is undermining criminal trials. In 2009, a National Academy of Sciences report identified a "notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods" and noted that numerous forensic disciplines lack known accuracy measures or error rates. Wexler, *supra* note 28, at 1421.

51. Aziz Z. Huq, *Racial Equity in Algorithmic Criminal Justice,* 68 Duke L. J. 1043, 1088 (2019). [hereinafter *Racial Equity in Algorithmic Criminal Justice].*

52. "Discriminatory intent" is a central term in the judicial interpretation of constitutional clauses requiring the equal treatment of persons without regard to their race, ethnicity, or religion. Aziz Z. Huq, *What Is Discriminatory Intent*, 103 Cornell L. Rev. 1211, 1212 (2018).

53. "(I)ncreasing evidence suggests that human prejudices have been baked into these tools because the machine-learning models are trained on biased police data. Far from avoiding racism, they may simply be better at hiding it. Many critics now view these tools as a form of techwashing, where a veneer of objectivity covers mechanisms that perpetuate inequities in society." *See* Will D. Heaven, *Predictive Policing Algorithms Are Racist. They Need to Be Dismantled*, MIT Tech. Rev., (July 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/.

54. *Racial Equity in Algorithmic Criminal Justice, supra* note 51, at 1076.

55. *The Undue Influence of Surveillance Technology Companies on Policing supra* note 48, at 132.

56. The Task Force defines "data-driven policing" as including, but not limited to, the surveillance technologies, tools, and methods employed by law enforcement to visualize crime; target "at-risk" individuals and groups; map physical locations; track digital communications; and collect data on individuals as well the communities they patrol. "Data-driven policing" also encompasses place-based predictive models that rely on historical crime data, geographic data, and demographic data; person-based predictive models that rely on personal data and social network analysis; and any databases, lists, and systems that subject individuals to increased police surveillance and monitoring.

57. Brief of Amici Curiae Public Justice Center, American Civil Liberties Union of Maryland, and Washington Lawyers' Committee for Civil Rights and Urban Affairs, *Sizer v. Maryland* (Md. Ct. App.) 9 (2017) (*quoting* Kelly Koss, *Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World*, 90 Chi-Kent L. Review 301, 312 (2015).

58. Wexler, *supra* note 28, at 1368.

59. Wexler, *supra* note 28, at 1429.

60. When new surveillance technologies are kept secret because of non-disclosure agreements, they cannot be challenged by criminal defendants and these challenges cannot be decided by judges, regardless of the merits of the defendants' claims. The use of a new surveillance technology may or may not be considered a Fourth Amendment search, but a private company's insistence on secrecy removes the legal issue from judicial review.

61. Laura M. Moy, *A Taxonomy of Police Technology's Racial Inequity Problems*, 2021 U. Ill. L. Rev. 139 (2021).

62. *Tracked and Trapped, supra* note 35.

63. Mariam Arain et al., *Maturation of the Adolescent Brain*, 9 Neuropsychiatric Disease and Treatment 449, 453 (2013).

64. Our flagship software, the Palantir Platform, features a full suite of analytical tools that enable organizations and their analysts to collaboratively generate actionable insights from large and disparate data sets. The Palantir Platform has been used to tackle data-driven problems in myriad contexts, from efficiently delivering aid to victims of a natural disaster to helping law enforcement agencies (LEAs) coordinate efforts to track a dangerous fugitive or rescue an abducted child." *See* Palantir Technologies, *Palantir and Law Enforcement: Protecting Privacy and Civil Liberties*, https://blog.predpol.com/geolitica-a-new-name-a-new-focus.

65. "The NYPD's precision policing consists of two important components: an intelligence-led investigative component and a neighborhood coordination component. The intelligence-led investigative side of this policy uses predictive analysis of all crimes to identify the small population of criminals who commit most of the violent crimes. [1] It also identifies small areas on the map with a higher intensity of crime, otherwise known as hot spots. [2] This new approach targets illegal firearms, precisely identifies gangs, and establishes a database for recidivists. The second component connects citizens with the police through an initiative known as neighborhood coordination. This program has three core goals: to reduce crime further, to promote trust and respect, and to solve problems by collaborating with residents. Many small community concerns are being solved collectively rather than through strict enforcement of minor crimes." *See* Muhammad Ashraf, Precision Policing: A Way Forward to Reduce Crime (2020), https://www.hsaj.org/articles/16249.

66. "Our original name of PredPol was a mashup of "predictive policing," and indeed we helped popularize and define that concept when we launched the company in 2012. The idea was that by analyzing historical data we could help better position patrol officers to prevent crime before it occurs. But that phrase has broadened to include activities – such as facial recognition or "predicting" that certain individuals will commit crimes – with which we are not aligned; even the use of the word "predictive" itself does not accurately describe our business." *See Geolitica: A New Name, A New Focus,* PredPol: Predictive Policing Technology, https://blog.predpol.com/geolitica-a-new-name-a-new-focus, (last visited Mar. 32, 2021).

67. Amna Akbar, *Toward a Radical Imagination of Law,* 3 N.Y.U. L. Rev. 93 (2018).

68. The mission of NACDL's Task Force on Predictive Policing is to study the issues surrounding the use of various data collection tools and analytical techniques that purport to prospectively identify where criminal activity is likely to occur and the people likely to be involved. The Task Force will evaluate the impact of these techniques on privacy and other individual constitutional rights, issue recommendations for best practices to safeguard those rights, and provide legal and technical assistance to educate defense practitioners in addressing the use of these tools and techniques. To arrive at these recommendations, the Task Force will focus on the stated goals of the various technologies being used, evidence of their effectiveness, and their potential shortcomings. The overarching principle driving the Task Force's evaluation and recommendations will be the implications of predictive policing on how individuals are surveilled, investigated, charged, and prosecuted. The Task Force will specifically address the implications of predictive policing for racial profiling and the impact of policing and prosecution on historically overpoliced groups and communities of color.

69. Every person who talked to the Task Force was informed that they could go off the record at any time. Some witnesses did go off the record at times, which is indicated in the transcripts. Two witnesses spoke with the Task Force on the condition that the conversation be completely off the record.

70. James Baldwin, "Black English: A Dishonest Argument", 1980, as quoted in *I Am Not Your Negro.*

71. Modern day policing emerged from the shadows of slave patrols and night-watch systems in the 1800s. As a profession, policing began in the mid-1800s. Boston is credited with forming the first full-time police force in the United States in 1838. Other states and cities soon followed Boston's lead and formed their own police departments. Sally Hadden, *Slave Patrols: Law and Violence in Virginia and the Carolinas* (2003).

72. *Id.*

73. Michelle Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (The New Press 2010).

74. Douglas A. Blackmon, *Slavery by Another Name,* 42 (2009).

75. In addition to the Ku Klux Klan, other paramilitary groups that were formed or existed during this period colluded with local politicians towards the same objective as Black Codes. *See* Eric Foner; *Black Reconstruction: An Introduction*, 112 South Atlantic Q. 409 (2013), *available at* https://doi.org/10.1215/00382876-2146368.

76. David M. Oshinsky, Worse Than Slavery: Parchman Farm and the Ordeal of Jim Crow Justice 21 (1996).

77. *Id.* at 21.

78. *Id.*

79. Nikole Hannah-Jones, "What Is Owed," *The New York Times Magazine,* June 28, 2020, at p. 50.

80. Wesley M. Oliver, The Prohibition Era and Policing: A Legacy of Misregulation (2018).

81. *See* Mapp v. Ohio, 367 U.S. 643 (1961).

82. *See* Gideon v. Wainwright, 372 U.S. 335 (1963).

83. *See* Miranda v. Arizona, 384 U.S. 436 (1966).

84. Christopher Klien, *How Selma's Bloody Sunday Became a Turning Point in the Civil Rights Movement, History*, History.com (July 18, 2020), https://www.history.com/news/selma-bloody-sunday-attack-civil-rights-movement.

85. Natsu Taylor Saito, *Whose Liberty? Whose Security? The USA PATRIOT Act in the Context of COINTELPRO and the Unlawful Repression of Political Dissent*, 81 Or. L. Rev. 1051, 1079-80 (2002).

86. Zero tolerance policing is sometimes known as "aggressive policing" or "aggressive order maintenance" and is sometimes incorrectly tied to "broken windows" policing. A zero tolerance strategy consists of stopping, questioning, and frisking pedestrians or drivers considered to be acting suspiciously and then arresting them for offenses when possible, typically for such low-level offenses as possessing marijuana. A defining difference between zero tolerance interventions and other strategies is that zero tolerance strategies are not discerning; the focus is on making stops and arrests to crack down on all types of disorder, generically defined. RAND Corp., *Zero Tolerance and Aggressive Policing (and Why to Avoid It) in Depth*, https://www.rand.org/pubs/tools/TL261/better-policing-toolkit/all-strategies/zero-tolerance/in-depth.html (last visited Apr. 6, 2021).

87. The broken windows model of policing was first described in 1982 in a seminal article by Wilson and Kelling. Briefly, the model focuses on the importance of disorder (e.g., broken windows) in generating and sustaining more serious crime. Disorder is not directly linked to serious crime; instead, disorder leads to increased fear and withdrawal from residents, which then allows more serious crime to move in because of decreased levels of informal social control. Ctr. for Evidence Based Crime Policy, *Broken Windows Policing*, https://cebcp.org/evidence-based-policing/what-works-in-policing/research-evidence-review/broken-windows-policing/ (last visited Apr. 6, 2021).

88. Under Lyndon B. Johnson's presidency, the Administration first began giving military equipment to law enforcement agencies. *See* Elizabeth Hinton, *A War Within Our Own Boundaries: Lyndon Johnson's Great Society and the Rise of the Carceral State,* 102 J. Am. Hist. 100 (2015), *available at* https://academic.oup.com/jah/article/102/1/100/686903.

89. The beginning of modern war on drugs in the United States is commonly credited to President Richard Nixon, who evoked fears of crime, degenerate youth, and foreign drugs to garner support for his massive, by early 1970s standards, effort to combat drugs in the United States. Anne L. Foster, *The Long War on Drugs*, *in* Oxford Res. Encyclopedias of Am. Hist. (2017), *available at* https://doi.org/10.1093/acrefore/9780199329175.013.402.

90. *Fusion Centers*, Dep't of Homeland Sec. (Sept. 19, 2019), HTTPS://WWW.DHS.GOV/FUSION-CENTERS.

91. Created as part of 1997's National Defense Authorization Act, the 1033 program allows the Department of Defense to get rid of excess equipment by passing it off to local authorities, who only have to pay for the cost of shipping. (A precursor, the slightly more restrictive 1208 program, began in 1990.) According to the Law Enforcement

Support Office (LESO), which oversees the process, over $7.4 billion of property has been transferred since the program's inception; more than 8,000 law enforcement agencies have enrolled. Brian Barnett, *The Pentagon's Hand-Me-Downs Helped Militarize Police. Here's How*, Wired (June 6, 2020, 4:54 PM), https://www.wired.com/story/pentagon-hand-me-downs-militarize-police-1033-program/.

92. *Oversight Hearing on Policing Practices and Law Enforcement Accountability Before the H. Comm. On the Judiciary*, 116th Cong. 1 (2020) (statement of Paul Butler, Albert Brick Professor in Law, Georgetown University Law Center), *available at* https://docs.house.gov/meetings/JU/JU00/20200610/110775/HHRG-116-JU00-Wstate-ButlerP-20200610.pdf.

93. *See, e.g.,* Floyd v. City of New York, 302 F.R.D. 69 (S.D.N.Y. 2014), *aff'd in part, appeal dismissed in part*, 770 F.3d 1051 (2d Cir. 2014); Illinois v. Wardlow, 528 U.S. 119 (2000); Malley v. Briggs, 457 U.S. 335 (1986).

94. *The New Surveillance Discretion supra* note 8, at 10.

95. Rashida Richardson, Jason Schultz, & Kate Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice* 94 N.Y.U. L. Rev. Online 192 (2019).

96. Wayne A. Logan & Andrew Guthrie Ferguson, *Policing Criminal Justice Data*, 101 Minn. L. Rev. 541 (2016).

97. Act to Establish the Department of Justice, ch. 150 § 12, 16 Stat. 162, 164.

98. Jed Handelsman Shugerman, *The Creation of the Department of Justice*: *Professionalization Without Civil Rights or Civil Service*, 66 Stan. L. Rev. 121 (2014).

99. U.S. Congress, Acts and Resolutions of the United States of America Passed at the Second Session of the Forty-First Congress, 1870.

100. J. L. Thompson, *Uniform Crime Reporting: Historical IACP Landmark*, 35 Police Chief 22, 23 (1968).

101. Following the 1890 census, crime data was used to impute color to crime, solidifying the link between race and criminality and justifying much of the racialized policing practices that came afterward. *See* Khalil Gibran Muhammad, The Condemnation of Blackness: Race, Crime, and the Making of Modern Urban America (2019).

102. Uniform Crime Reporting, FBI, https://ucr.fbi.gov/ (last visited June 17, 2021); Clayton Mosher et al., The Mismeasure of Crime 43 (2d ed. 2011).

103. John Koren, *Report of Committee on Statistics of Crime*, 1 J. Crim. L. & Criminology 417 (1910).

104. A. Vollmer, *Criminal Statistics*, in IACP, op. cit. supra note 100, p. 72.

105. William Samuel Isaac, *Hope, Hype, and Fear: The Promise and Potential Pitfalls of the Big Data Era in Criminal Justice*, 15 Ohio St. J. of Crim. L. 1, 2 (2018).

106. Donald J. Black, *Production of Crime Rates*, 35 Am. Socio. Rev. 733, 735 (1970).

107. *Id.*

108. *The New Surveillance Discretion*, supra note 8, at 14.

109. *Id.* at 9.

110. Isaac, *supra* note 105, at 2.

111. Ronald H. Beattie, *The Sources of Criminal Statistics,* 217 Annals of the Am. Acad. of Pol. & Soc. Sci. 19, 21–22 (1941).

112. Black, *supra* note 106, at 734.

113. *The New Surveillance Discretion*, *supra* note 8, at 10.

114. Task Force Interview with Jay Stanley at 92, Senior Policy Analyst. ACLU Speech, Privacy, and Technology Project, in Washington, D.C. (Dec. 18, 2017).

115. John A. Eterno, Arvind Verma & Eli B. Silverman, *Police Manipulations of Crime Reporting: Insiders' Revelations*, 33 Just. Q. 811 (2016).

116. *Id.*

117. Logan & Ferguson, *supra* note 96, at 550.

118. Stephen D. Mastrofski & James J. Willis, *Police Organization Continuity and Change: Into the Twenty- first Century,* 39 Crime and Just. 55, 87 (2010).

119. *Id.* at 88.

120. *Id.*

121. Rashida Richardson & Amba Kak, *It's Time for a Reckoning About this Foundational Piece of Police Technology*, Slate (Sept. 11, 2020, 1:38 PM), https://slate.com/technology/2020/09/its-time-for-a-reckoning-about-criminal-intelligence-databases.html.

122. Logan & Ferguson, *supra* note 96, at 543.

123. *Herring v United States*, 555 U.S. 135 (2009).

124. Logan & Ferguson, *supra* note 96, at 543.

125. K.A. Taipale, *Data Mining and Domestic Security: Connecting the Dots to Make Sense of Data*, 5 Ctr. Advanced Stud. Sci. Tech. Poly 1, 21 (2004).

126. "Query" is defined as the search of a database for all records satisfying some specified condition. *Id.* at 22.

127. *Id.* at 13.

128. Size refers to the number of records or objects in the database, and dimensionality refers to the number of fields or attributes to an object.

129. Taipale, *supra* note 125, at 29.

130. *Id.*

131. *Id.* at 22.

132. An algorithm is generally regarded as the mathematical logic behind any type of system that performs tasks or makes decisions. *See* AI Now Institute, Algorithmic Accountability Policy Toolkit 2 (October 2018).

133. Solon Barocas et al., Data and Soc'y Res. Inst., Data & Civil Rights Technology Primer 1, 3 (2014).

134. Task Force Interview with Daniel Kahn Gillmor at 122, Senior Staff Technologist. ACLU Speech, Privacy, and Technology Project, in Washington, D.C. (Dec. 18, 2017).

135. Andrew D. Selbst, Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, & Janet Vertesi, *Fairness and Abstraction in Sociotechnical Systems*, in Proceedings of the Conference on Fairness, Accountability, and Transparency, Ass'n for Computing Machinery (2019), *available at* https://doi.org/10.1145/3287560.3287598.

136. Taipale, *supra* note 123, at 21.

137. *Id.*

138. *Big Data and Predictive Reasonable Suspicion, supra* note 10, at 330.

139. Aaron Rieke et al., Open Soc'y Found., Data Brokers in an Open Society 35 (2016).

140. *Id.*

141. Richardson & Kak, *supra* note 121.

142. Rieke et al., *supra* note 139, at 38.

143. Daniel J. Steinbock, *Data Matching, Data Mining, and Due Process*, 40 Ga. L. Rev. 1, 14 (2005).

144. Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 Calif. L. Rev. 671, 677 (2016).

145. *Id.*

146. Isaac, *supra* note 105, at 4.

147. Taipale, *supra* note 125, at 13.

148. For the purposes of this report, GIS technologies can encompass any computer-based tools used to "modify, visualize, query, and analyze geographic and tabular data." GIS also includes the "development of particular software programs that help researchers visualize data, assess human behavior over geographic space, follow spatial patterns, validate theories, and examine how geography affects crime and public safety." *See* Andrew Guthrie Ferguson, *Crime Mapping and the Fourth Amendment: Redrawing High-Crime Areas,* 63 Hastings L.J. 179, 184 (2011) [hereinafter Crime Mapping and the Fourth Amendment].

149. *Id.* at 184.

150. *Id.* at 108.

151. Paul and Patricia Brantingham are related to Jeff Brantingham, one of the co-founders of the predictive policing program "Geolitica" and an anthropology professor at the University of California, Los Angeles.

152. Paul Brantingham & Patricia Brantingham, Environmental Criminology (Sage Publications 1981).

153. *Id.* at 69.

154. Environmental Criminology and Crime Analysis 12 (Richard Wortley & Lorraine Mazerolle eds., 2008).

155. *Id.* at 99.

156. Ned Levine, *Crime Mapping and the Crimestat Program*, 38 Geographical Analysis, 41, 42 (2005).

157. *Environmental Criminology and Crime Analysis supra* note 154, at 100.

158. *Crime Mapping and the Fourth Amendment supra* note 148, at 104.

159. *Id.* at 110.

160. Richardson *supra* note 95.

161. *Environmental Criminology and Crime Analysis supra* note 150, at 13.

162. Wortley & Mazerolle, *supra* note 154, at 101.

163. David Weisburd et al., *Reforming to Preserve: Compstat and Strategic Problem Solving in American Policing*, 2 Criminology Pub. Pol'y 421, 424 (2003).

164. U.S. Dep't. of Just., Bureau of Just. Assistance, CompStat: Its Origins, Evolution and Future in Law Enforcement Agencies 3 (2013), https://bja.ojp.gov/sites/g/files/xyckuh186/files/Publications/PERF-Compstat.pdf.

165. *Policing By Numbers, supra* note 6, at 43.

166. William K. Rashbaum, *Retired Officers Raise Questions on Crime Data*, N.Y. Times (Feb. 6, 2010), https://www.nytimes.com/2010/02/07/nyregion/07crime.html.

167. *1990s Drop in NYC Crime Not Due to CompStat, Misdemeanor Arrests, Study Finds*, N.Y.U. (Feb. 4, 2013), https://www.nyu.edu/about/news-publications/news/2013/february/1990s-drop-in-nyc-crime-not-due-to-compstat-misdemeanor-arrests-study-finds.html.

168. Isaac, *supra* note 105, at 4.

169. Ingrid Burrington, *What Amazon Taught Cops: Predictive Policing is Just Another Form of Supply-Chain Efficiency,* The Nation (May 27, 2015), https://www.thenation.com/article/archive/what-amazon-taught-cops/.

170. Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 82 Am. Socio. Rev. 977, 981 (2017).

171. PredPol has since been re-branded as "Geolitica."

172. Lum & Isaac, *supra* note 18, at 14.

173. *Policing Predictive Policing, supra* note 15, at 1127.

174. Caroline Haskins, *Revealed: This is Palantir's Top-Secret User Manual for Cops*, Vice (July 12, 2019), https://www.vice.com/en/article/9kx4z8/revealed-this-is-palantirs-top-secret-user-manual-for-cops . *(Editor's Note: This report originally cited pages from Palantir's website detailing the organization's work with police departments across the United States. By Oct. 2021, however, Palantir removed any documentation of its efforts to assist law enforcement from its website. This citation has been adjusted to reflect that change.)*

175. Burrington, *supra* note 169.

176. Michael Kwet, *The Microsoft Police State: Mass Surveillance, Facial Recognition and the Azure Cloud*, The Intercept (June 14, 2020, 3:42 PM), https://theintercept.com/2020/07/14/microsoft-police-state-mass-surveillance-facial-recognition/.

177. *Records and Evidence Management*, Motorola, https://www.shotspotter.com/law-enforcement/patrol-management/ (last visited Mar. 31, 2021).

178. *Crime Analytics and Mapping*, Lexis Nexis Risk Solutions https://risk.lexisnexis.com/law-enforcement-and-public-safety/crime-analytics-and-mapping (last visited Mar. 31, 2021).

179. *Community First Patrol Management Software Improves Crime Deterrence: Mitigates Over Policing and Biased Patrols for Positive Community Engagement*, Shotspotter (last visited Mar. 31, 2021), https://www.shotspotter.com/law-enforcement/patrol-management/.

180. *Feeding the Machine, supra* note 24, at 302.

181. *Policing By Numbers, supra* note 6, at 42

182. *Policing Predictive Policing, supra* note 15, at 1132.

183. *Policing By Numbers, supra* note 6, at 44.

184. *Policing Predictive Policing, supra* note 15, at 1123.

185. *Id.* at 1125.

186. *Feeding the Machine, supra* note 24, at 5.

187. Richardson, *supra* note 95, at 198.

188. *Id.*

189. Julie Barrows & C. Ronald Huff, *Gangs and Public Policy-Constructing and Deconstructing Gang Databases,* 8 Criminology & Pub. Pol'y 675, 679 (2009).

190. Megan Behrman, Note, *When Gangs Go Viral: Using Media and Surveillance Cameras to Enhance Gang Databases,* Harv. J. L. & Tech. 315, 316 (2015).

191. "In California, as of 2016, men of color made up 90,000 out of the 106,000 individuals, or about 93.4%, on the statewide database, CalGang. In New York City, the NYPD's database is almost entirely comprised of people of color, with only 1.1% of white individuals listed. Of the individuals alleged to be 'gang-affiliated' on Boston's database, 66% were Black, 24% were Latinx, and just 2% were white. The use of gang databases can therefore increase the disproportionate use of raids and criminal prosecutions against people of color." Sofia Lopez-Franco, *Criminalizing Gang Databases: Why People of Color Need to Go*, RaceNYU (Jan. 28, 2020) https://medium.com/@RaceNYU/criminalizing-people-of-color-why-gang-databases-need-to-go-d8b04645b7c5.

192. K. Babe Howell, *Gang Policing: The Post Stop-and-Frisk Justification for Profile-Based Policing,* 5 Den. Crim.

L. Rev. 1, 15 (2015).

193. *Id.*

194. *Social Media Fact Sheet*, Pew Res. Ctr. (Apr. 7, 2021), https://www.pewresearch.org/internet/fact-sheet/social-media/.

195. In her meeting with the Task Force, Rachel Levinson-Waldman, Senior Counsel to the Liberty and National Security Program at the Brennan Center for Justice, explained that some third-party companies had advertised to law enforcement "their ability to track protestors." She stated that in response, Twitter, Facebook, and Instagram "changed their terms of service . . . to say you can't use our date for surveillance purposes." Task Force Interview with Rachel Levinson-Waldman at 72, Senior Counsel to the Liberty and National Security Program, Brennan Center for Justice, in Washington, D.C., (Dec. 18, 2017).

196. Desmond Upton Patton, *Stop and Frisk Online: Theorizing Everyday Racism in Digital Policing in the Use of Social Media for Identification of Criminal Conduct and Associations*, 2017 Social Media + Soc'y 1, 2.

197. *Id.* at 3.

198. *Id.*

199. *See, e.g.*, Natasha Duarte, Emma Llanso, & Anna Loup, *Mixed Messages? The Limits of Automated Social Media Content Analysis*, Ctr. for Democracy & Tech. (Nov. 2017), http://proceedings.mlr.press/v81/duarte18a.html.

200. *Bringing Big Data to the Enterprise*, IBM, https://web.archive.org/web/20170530065140/http://www-01.ibm.com:80/software/data/bigdata/.

201. Jim Gray & Shenoy Prashant, Microsoft Res. Advanced Tech. Div., Rules of Thumb in Data Engineering (2009), http://research.microsoft.com/pubs/68636/ms_tr_99_100_rules_of_thumb_in_data_%20engineering.pdf.

202. Marcus Wohlsen, *Dropbox Slashes Its Price as The Cost of a Gigabyte Nears Zero*, Wired (Aug. 7 2014, 9:00 AM) https://www.wired.com/2014/08/dropboxs-plan-to-stay-relevant/.

203. *Policing By Numbers, supra* note 6, at 40.

204. Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. Rev. 93, 96 (2014).

205. Brayne, *supra* note 170, at 980.

206. *Id.*

207. *Big Data and Predictive Reasonable Suspicion, supra* note 10.

208. Perry, *supra* note 12.

209. Barocas et al., *supra* note 133, at 4.

210. *Id.*

211. *Id.*

212. *Predictive Policing Theory, supra* note 13.

213. Barrett, *supra* note 11, at 334.

214. *Policing Predictive Policing, supra* note 15 at 1113.

215. *Id.*

216. AI Now Institute, *supra* note 132, at 2.

217. In his meeting with the Task Force in 2019, Philip Stark, a Professor of Statistics at the University of

California, Berkeley, was critical of the use of the ETAS model in PredPol; claiming that the ETAS model itself is "not that great for [predicting earthquakes] in seismology" and that its use in the PredPol model demonstrates that "PredPol is trying to borrow strength from something that [itself] doesn't actually have strength, much less is a parallel to the occurrence of crime." *See* Task Force Interview with Philip Stark at 47, University of California, Berkeley (Jun. 27, 2019).

218. G. O. Mohler et al., *Randomized Controlled Field Trials of Predictive Policing,* J. Am. Stat. Ass'n, 1399 (2015).

219. *Geolitica: A New Name, A New Focus*, PredPol: Predictive Policing Technology Blog, HTTPS://BLOG.PREDPOL.COM/GEOLITICA-A-NEW-NAME-A-NEW-FOCUS (last visited June 17, 2021).

220. *Predictive Policing Technology*, PredPol, https://www.predpol.com/technology/ (last visited June 17, 2021).

221. David Speiglehalter, *Should We Trust Algorithms?*, 2 Harv. Data Sci. R. 1 (2020), https://hdsr.mitpress.mit.edu/pub/56lnenzj/release/1.

222. *Id.*

223. *PredPol's Stance on Privacy Civil Rights and Technology*, PredPol: Predictive Policing Blog, https://blog.predpol.com/predpols-stance-on-privacy-civil-rights-transparency, (last visited June 17, 2021).

224. *Predictive Policing Technology*, PredPol, HTTPS://WWW.PREDPOL.COM/TECHNOLOGY/ (last visited June 17, 2021).

225. Task Force Interview with Professor Andrew Ferguson, University of the District of Columbia, David A. Clarke School of Law, Washington D.C. (Dec. 8, 2017).

226. Perry, *supra* note 12.

227. *Id.* at 9.

228. *Id.*

229. Task Force Telephone Interview with Sean Malinowski at 28, Director of Policing Innovation and Reform at the University of Chicago Crime Lab, (Aug. 14, 2019).

230. Task Force Telephone Interview with Professor Jeffery Brantingham at 28, UCLA, (June 27, 2019).

231. PredPol, *supra* note 220.

232. Perry, *supra* note 12, at 41.

233. Jeff Brantingham, PredPol, *Predictive Policing in Action* 8 (2012) https://www.columbiasc.net/depts/city-council/docs/old_downloads/07_17_2012_Agenda_Items/PredPol%20One%20Pager%20Columbia%20Richland%20County%202012%20June.pdf.

234. *Policing Predictive Policing, supra* note 15, at 1132.

235. As a point of comparison, Risk Terrain Modeling (RTM) is a strategy that specifically examines the risk environment of an area to figure out why crimes are occurring by seeing the "physical reality of a city as a terrain of overlapping risks," with the more risks in close proximity to each other leading to a heightened risk of forecast crime. Unlike PredPol, the goal of RTM is to look for fixed place-based factors as inputs to create a risk narrative for particular crimes, under the theory that because spatial environments can encourage criminal behavior, fixing those same spaces can reduce criminal acts. *See Policing Predictive Policing supra* note 15, at 1132.

236. Task Force Telephone Interview with Sean Malinowski at 22, Director of Policing Innovation and Reform at the University of Chicago Crime Lab (Aug. 14, 2019).

237. Task Force Interview with Zach Friend at 13-14, Second District Supervisor for Santa Cruz County and former Press Information Officer and Crime Analyst for the Santa Cruz, Ca. Police Department, in L.A. (Feb. 8, 2018).

238. *Id.*

239. *Id.*

240. Caroline Haskins, *Dozens of Cities Have Secretly Experimented With Predictive Policing Software*, Vice (Feb. 6, 2019, 10:00 AM), https://www.vice.com/en/article/d3m7jq/dozens-of-cities-have-secretly-experimented-with-predictive-policing-software.

241. Rachael Myrow, *Are California Police Departments Slowly Backing Away from Predictive Policing?* KQED (June 12, 2020) https://www.kqed.org/news/11828568/a-technical-fix-may-not-help-address-systematic-racism-in-predictive-policing.

242. Mark Puente, *LAPD Pioneered Predicting Crime with Data. Many Police Don't Think It Works*, L.A. Times (July 3 2019), https://www.latimes.com/local/lanow/la-me-lapd-precision-policing-data-20190703-story.html.

243. Task Force Interview with Jeremy Heffner at 216-17, Former Senior Data Scientist and Product Manager for Azavea's HunchLab, in Chicago, Ill. (June 7, 2018). Mr. Heffner has since left HunchLab.

244. Aaron Shapiro, *Predictive Policing for Reform? Indeterminacy and Intervention in Big Data Policing,* 17 Surveillance & Soc'y 456, 461 (2019).

245. Robert Cheetham, *Why We Sold HunchLab*, Azavea (Jan. 23, 2019), https://www.azavea.com/blog/2019/01/23/why-we-sold-hunchlab/.

246. *ShotSpotter Announces Acquisition of HunchLab to Springboard into AI Driven Analysis and Predictive Policing*, ShotSpotter (Oct. 3, 2018), https://www.shotspotter.com/press-releases/shotspotter-announces-acquisition-of-hunchlab-to-springboard-into-ai-driven-analysis-and-predictive-policing/.

247. Azavea, *HunchLab: Under the Hood* (2015), https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf

248. Shapiro, *supra* note 244, at 462.

249. Azavea, Answers to Questionnaire: Hunchlab 5, https://www.brennancenter.org/sites/default/files/NYC_0000679_HunchlabAnswerstoQuestionnaire%20-%20Copy.pdf

250. A 2017 U.S. Commission Report that found increased judicial discretion leads to greater racial disparity in the federal criminal system, with an established "gap between the sentence lengths for Black and White male offenders." USSC, *Demographic Differences in Sentencing: An Update to the 2012 Booker Report* at 2 (2017).

251. *Id.*

252. Task Force Interview with Jeremy Heffner at 216-17, Former Senior Data Scientist and Product Manager for Azavea's *HunchLab,* in Chicago, Ill. (June 7, 2018).

253. *Id.*

254. *Id.*

255. *Id.*

256. Cheetham, *supra* note 245.

257. Tim Lau, *Predictive Policing Explained*, Brennan Ctr. for Just. (Apr. 1, 2020), https://brennancenter.org/our-work/research-reports/predictive-policing-explained

258. Aaron Tucek, *Constraining Big Brother: The Legal Deficiencies Surrounding Chicago's Use of the Strategic Subject List,* 2018 U. Chi. Legal F. 427, 432.

259. Matt Stroud, *This 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.

260. *Id.*

261. Custom Notifications in Chicago, Spec. Order S10-05 (Chi. Police Dep't Oct. 6, 2015), http://directives.chi-cagopolice.org/directives/data/a7a57bf0-1456faf9-bfa14-570a-a2deebf33c56ae59.html [perma.cc/2RA5-456U] [hereinafter SPEC. ORDER S10-05]. (*Editor's Note: As of Oct. 13, 2021, this link is no longer available to the public.)*

262. Briana Posadas, *How Strategic Is Chicago's "Strategic Subjects List"? Upturn Investigates*, Medium (June 22, 2017), https://medium.com/equal-future/how-strategic-is-chicagos-strategic-subjects-list-upturn-investigates-9e5b4b235a7c.

263. Tucek, *supra* note 258, at 432.

264. *Id.*

265. *Id.*

266. Task Force Interview with David Robinson at 195, Managing Director and Co-Founder of Upturn, in Washington D.C. (Dec. 18, 2017).

267. Tucek, *supra* note 258, at 434.

268. Jessica Saunders, Priscilla Hunt, & John S. Hollywood, *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. Experimental Criminology 347 (2016).

269. *Id.*

270. Andrew V. Papachristos & Christopher Wildeman, *Network Exposure and Homicide Victimization in an African American Community,* 104 Am. J. Pub. Health 143 (2014).

271. Stroud, *supra* note 259.

272. *Tracked and Targeted, supra* note 40.

273. Spec. Order S10-05, *supra* note 261.

274. Josh Kaplan, *Predictive Policing and the Long Road to Transparency*, The Southside Wkly. (July 12, 2017), HTTPS://SOUTHSIDEWEEKLY.COM/PREDICTIVE-POLICING-LONG-ROAD-TRANSPARENCY/.

275. Posadas, *supra* note 262.

276. Task Force Interview with Freddy Martinez at 103, Director, Lucy Parsons Labs, in Chicago, Ill. (June 7, 2018). Mr. Martinez appeared telephonically.

277. Jeff Asher & Rob Arthur, *Inside the Algorithm that Tries to Predict Gun Violence in Chicago,* N.Y. Times (June 13, 2017), https://www.nytimes.com/2017/06/13/upshot/what-an-algorithm-reveals-about-life-on-chicagos-high-risk-list.html.

278. Andrew V. Papachristos, *CPD's Crucial Choice: Treat Its List as Offenders or as Potential Victims?*, Chi. Trib. (July 26, 2016, 10:00 AM), https://www.chicagotribune.com/opinion/commentary/ct-gun-violence-list-chicago-police-murder-perspec-0801-jm-20160729-story.html.

279. *Id.*

280. Task Force Interview with Chaclyn Hunt at 104, Director of the Youth / Police Project, Invisible Institute, in Chicago, Ill. (June 7, 2018).

281. Yana Kunichoff & Patrick Sier, *The Contradictions of Chicago Police's Secretive List,* Chi. Mag. (Aug. 21, 2017, 8:44 AM), https://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/.

282. *Tracked and Targeted, supra* note 40, at 5.

283. Mick Dumke & Frank Main, *A Look Inside the Watch List Chicago Police Fought to Keep Secret*, Chi. Sun-Times (May 18, 2017, 9:26 AM), https://chicago.suntimes.com/2017/5/18/18386116/a-look-inside-the-watch-list-chicago-police-fought-to-keep-secret.

284. Posadas, *supra* note 262.

285. Jeremy Gorner & Annie Sweeney, *For Years Chicago Police Rated the Risk of Tens of Thousands Being Caught up in Violence. That Controversial Effort has Quietly Been Ended,* Chi. Trib. (Jan. 24, 2020, 8:55 PM) https://www.chicagotribune.com/news/criminal-justice/ct-chicago-police-strategic-subject-list-ended-20200125-spn4kjmrxrh4tmktdjckhtox4i-story.html.

286. Aaron Stagoff-Belfort, *The Lessons of Chicago's Disastrous "Crime Prediction" Experiment*, FILTER (Mar. 12, 2020), https://filtermag.org/chicago-crime-prediction.

287. Task Force Interview with Jessica Saunders at 60, Criminologist at RAND, in L.A. (Feb. 9, 2018). (Ms. Saunders has since left her position at RAND).

288. Spec. Order S10-05, *supra* note 261.

289. Luke Munn, Seeing with Software: Palantir and the Regulation of Life, 2 Stud. in Control Societies 1 (2017).

290. Caroline Haskins, *Revealed: This Is Palantir's Top-Secret User Manual for Cops, Tech*, Vice (July 12, 2019), https://www.vice.com/en/article/9kx4z8/revealed-this-is-palantirs-top-secret-user-manual-for-cops/ [hereinafter *Revealed: This Is Palantir's Top-Secret User Manual for Cops*].

291. Munn, *supra* note 289.

292. *Revealed: This Is Palantir's Top-Secret User Manual for Cops, supra n*ote 290.

293. Ali Winston, *Palantir Has Secretly Been Using New Orleans to Test Its Predicative Policing Technology*, The Verge (Feb 27, 2018, 3:25 PM), https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd.

294. Emily Lane, *Mayor, Police Chief to Face Subpoenas from Convicted Gang Member over Palantir Claim*, NOLA.com (July 12, 2019, 12:03 PM), https://www.nola.com/news/crime_police/article_fa5949c4-a300-509d-90e8-2d7814f505f6.html.

295. *Id.*

296. Matt Sledge, *Convicted Gang Leader Can Challenge NOPD's Use of Crime-Fighting Software, Judge Rules,* NOLA.com (Mar. 14, 2018, 2:09 PM), https://www.nola.com/article_0bb6a63e-00cc-5f2e-a367-3f624877a76a.html.

297. Winston, *supra* note 293.

298. Lane, *supra* note 294.

299. Michael Isaac Stein, *Months After End of 'Predictive Policing' Contract, Cantrell Administration Works on New Tool to Id 'High-Risk" Residents*, The Lens (Oct. 4, 2018), https://thelensnola.org/2018/10/24/months-after-end-of-predictive-policing-contract-cantrell-administration-works-on-new-tool-to-id-high-risk-residents/.

300. Task Force Interview with Kevin Vogeltanz at 75, Attorney, in Washington D.C. (May 5, 2019) (Mr. Vogeltanz appeared via video teleconference).

301. Johnathan Bullington & Emily Lane, *How a Tech Firm Brought Data and Worry to New Orleans Crime Fighting*, Times-Picayune (Mar. 1, 2018, 10:47 AM), https://www.nola.com/news/crime_police/article_33b8bf05-722f-5163-9a0c-774aa69b6645.html.

302. Winston, *supra* note 293.

303. *Id.*

304. Vogeltanz, *supra* note 300.

305. Bullington & Lane, *supra* note 301.

306. *Id.*

307. *Id.*

308. Vogeltanz, *supra* note 300.

309. Eva Ruth Moravec, *Do Algorithms Have a Place in Policing*, The Atlantic (Sept. 5, 2019), https://www.theatlantic.com/politics/archive/2019/09/do-algorithms-have-place-policing/596851/.

310. Eva Ruth Moravec, *Do Algorithms Have a Place in Policing*, The Atlantic (Sept. 5, 2019), https://www.theatlantic.com/politics/archive/2019/09/do-algorithms-have-place-policing/596851/.

311. Brenda Gazzar, *Activists File Lawsuit Over LAPD's Predictive Policing Program*, Gov't Tech. (Feb. 14, 2018), https://www.govtech.com/public-safety/Activists-File-Lawsuit-Over-LAPDs-Predictive-Policing-Program.html.

312. Task Force Interview with Professor Sarah Brayne at 45-46, University of Texas, Austin, in Washington D.C. (Dec 18, 2017).

313. Mara Hvistendahl, *How the LAPD and Palantir use Data to Justify Racist Policing*, The Intercept (Jan. 30, 2021), (https://theintercept.com/2021/01/30/lapd-palantir-data-driven-policing/).

314. Brayne, *supra* note 170, at 997.

315. *Policing Predictive Policing, supra* note 15, at 1142.

316. Maha Ahmed *Aided by Palantir, the LAPD Uses Predictive Policing to Monitor Specific People and Neighborhoods*, The Intercept (May 11, 2018) https://theintercept.com/2018/05/11/predictive-policing-surveillance-los-angeles/.

317. Hvistendahl, *supra* note 313.

318. Brayne, *supra* note 170, at 997.

319. *Id.*

320. Task Force Interview with Jamie Garcia at 168, Stop LAPD Spying Coalition., in L.A. (Feb 8, 2018).

321. Craig Uchida & Marc L. Swatt, *Operation LASER and the Effectiveness of Hotspot Patrol: A Panel Analysis*, 16 Police Q. 287 (2013).

322. For example, in 2018, 524 hours were recorded in the Pacific Division LASER zones, while 53,841 hours were recorded in the LASER zones of the Hollenbeck Division; indicating that the LASER program had incorrectly counted time spent parked at LAPD facilities as "dosage," often rendering officers' activity logs unreliable and making it difficult to draw conclusions about the effectiveness of the system in reducing vehicle or other crime. *See* Martin Macias Jr., *Audit Finds LAPD Predictive Policing Programs Lack Oversight*, Courthouse News (Mar. 8, 2019), https://www.courthousenews.com/audit-finds-lapd-predictive-policing-programs-lack-oversight/.

323. Mark Puente, *LAPD Ends Another Data-Driven Crime Program Touted to Target Violent Offenders*, L.A. Times, (Apr. 12, 2019, 4:48 PM), https://www.latimes.com/local/lanow/la-me-laser-lapd-crime-data-program-20190412-story.html.

324. *Id.*

325. James Blum, *The NYPD's Gang Database: A New Age of Stop and Frisk*, Stop Surveillance Tech. Oversight Project (July 23, 2019), https://www.stopspying.org/latest-news/2019/7/23/the-nypds-gang-database-a-new-age-of-stop-and-frisk.

326. *Id.*

327. Rachel Levinson-Waldman & Angel Díaz, *How to Reform Police Monitoring of Social Media, Stream,* Brookings Institute (July 9, 2020), https://www.brookings.edu/techstream/how-to-reform-police-monitoring-of-social-media/.

328. William Alden, *There's a Fight Brewing Between the NYPD and Silicon Valley's Palantir*, Buzzfeed News (June 28 2017, 3:23 PM), https://www.buzzfeednews.com/article/williamalden/theres-a-fight-brewing-between-the-nypd-and-silicon-valley.

329. Emily Hockett & Michael Price, *Palantir Contract Dispute Exposes NYPD's Lack of Transparency*, Just Security (July 20, 2017), https://www.justsecurity.org/43397/palantir-contract-dispute-exposes-nypds-lack-transparency/.

330. Dan Quart, *Ending Cy Vance's Gang Database*, Medium (Sept. 9, 2019), https://medium.com/assemblymember-dan-quarts-policy-prescriptions-for/changing-the-approach-on-gang-prosecutions-7bcbf519efa5.

331. Alden, *supra* note 328.

332. Ben Popper, *How the NYPD is Using Social Media to Put Harlem Teens Behind Bars,* The Verge, (Dec. 10, 2014, 1:15PM), https://www.theverge.com/2014/12/10/7341077/nypd-harlem-crews-social-media-rikers-prison

333. Richard Esposito, *New York's Kelly Plans 'Crew Cut' for Gang Members*, ABC News (Oct. 1, 2012, 7:33 PM), https://abcnews.go.com/Blotter/nypd-plans-crew-cut-gang-members/story?id=17370903.

334. Rose Hackman, *Is Online Surveillance of Black Teenagers the New Stop and Frisk?*, The Guardian (Apr. 23, 2015), https://www.theguardian.com/us-news/2015/apr/23/online-surveillance-black-teenagers-new-stop-and-frisk.

335. Kristine Conley, *Throw Book at 'Em*, N.Y. Post (Dec. 6, 2012, 5:00 AM), https://nypost.com/2012/12/06/throw-book-at-em/.

336. Task Force Interview with Jarrell Daniels at 81, Open Society Youth Activist Fellow and Founder of the Justice Ambassadors Youth Council Program at Columbia University's Center for Justice, in Washington D.C. (May 14, 2019). (Mr. Daniels appeared via teleconference).

337. *Id.*

338. *Id.*

339. Olivia Heffernan, '*We've Got One in the Sweep,*' The Appeal (July 30, 2019), https://theappeal.org/weve-got-one-in-the-sweep/.

340. Hackman, *supra* note 334.

341. Alice Speri, *The Largest Gang Raid in NYC History Swept up Dozens of Young People Weren't in Gangs*, The Intercept (Apr. 25, 2019), https://theintercept.com/2019/04/25/bronx-120-report-mass-gang-prosecution-rico/ [hereinafter *The Largest Gang Raid in NYC History Swept up Dozens of Young People Who Weren't in Gangs*].

342. *Id.*

343. Howell, *supra* note 192, at 21.

344. Speri, *supra* note 341.

345. Task Force Interview with Josmar Trujillo at 114, Activist and Writer, in N.Y.C (Apr. 18, 2018).

346. Julie Barrows & C. Ronald Huff, *Gangs and Public Policy-Constructing and Deconstructing Gang Databases,* 8 Criminology and Pub. Pol'y 675, 679 (2009).

347. Muniz, *supra* note 33.

348. *E.g.,* Hanna Dreier, *How a Crackdown on MS-13 Caught Up Innocent High School Students,* N.Y. Times Mag. (Dec. 27, 2018) (reporting on partnership in New York among ICE, local police departments and officers in public schools "to target and detain Latino immigrants suspected of gang ties").

349. Muniz, *supra* note 33.

350. *Id.*

351. *Id.*

352. N.Y. Crim. Proc. Law § 160.50.

353. N.Y. Crim. Proc. Law § 160.55.

354. Eli Hager, *Your Arrest was Dismissed. But It's Still in a Police Database*, Marshall Project (July 18, 2019, 4:00 AM), https://www.themarshallproject.org/2019/07/18/your-arrest-was-dismissed-but-it-s-still-in-a-police-database.

355. Brian Jefferson, Digitize and Punish, ch. 2 (2020), *available at* https://manifold.umn.edu/read/digitize-and-punish/section/c0ca86f2-5d33-4614-b74f-5ba837d2e981.

356. *Id.* at ch. 3.

357. Brief of Amici Curiae Pub. Justice Ctr., Am. Civil Liberties Union of Md., and Wash. Lawyers' Comm. for Civil Rights and Urban Affairs at 9, *Sizer v. State*, 174 A.3d 326 (Md. 2017) (No. COA-REG-0001-2017) (*quoting* Kelly Koss, *Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post*-Wardlow *World*, 90 Chi-Kent L. Rev. 301, 312 (2015)).

358. *Policing Predictive Policing, supra* note 15, at 1125.

359. *Policing By Numbers, supra* note 6, at 44.

360. Molly Griffard, *A Bias-Free Predictive Policing Tool?: An Evaluation of the NYPD's Patternizr*, 47 Fordham Urb. L.J. 43, 55 (2019).

361. *The New Surveillance Discretion, supra* note 8, at 18.

362. *Policing By Numbers, supra* note 6, at 58.

363. *Feeding the Machine, supra* note 24, at 302.

364. *The New Surveillance Discretion supra,* note 8, at 9.

365. Richardson, *supra* note 95, at 195.

366. *Stuck in a Pattern, supra* note 14.

367. *Feeding the Machine,* supra note 24, at 302.

368. Richardson, *supra* note 95, at 197.

369. *Id.* at 199.

370. *Feeding the Machine, supra* note 24, at 290.

371. *Id.*

372. *Id.*

373. *The New Surveillance Discretion, supra* note 8, at 18.

374. Lum & Isaac, *supra* note 18, at 17.

375. *Id.*

376. *Id.* at 16.

377. *Id.*

378. Barrett, *supra* note 11, at 337.

379. *Id.* at 341.

380. Jeffrey Brantingham and George Mohler later conducted their own study, which they say indicates their methods do not result in higher minority arrest rates. In 2018, they published a paper that began with acknowledging their critics, including Lum and Issac: "Though all of these studies deal with hypothetical scenarios or thought experiments, they succeed in demonstrating that careful attention needs to be paid to

whether predictive policing produces biased arrests." *See* Jeffrey Brantingham, Matthew Valasik & George O. Mohler, *Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial*, 5 Stat. & Pub. Pol'y 1, 2, (2018).

381. Though their 2016 study attracted criticism regarding the "appropriateness of using drug crime data for generating forecasts using PredPol's software, which has not been used for forecasting drug crimes," the authors argued that this critique ignores continued attempts by law enforcement to expand the scope of place-based predictive policing into drug crimes. *See* Kristian Lum & William Isaac, *Setting the Record Straight on Predictive Policing and Race*, Medium: In Justice Today (Jan. 3, 2018), https://medium.com/in-justice-today/setting-the-record-straight-on-predictive-policing-and-race-fe588b457ca2

382. Lum & Isaac, *supra* note 18, at 17.

383. *Id.* at 18.

384. *Id.*

385. Nil-Jana Akpinar, Maria De-Arteaga, & Alexandra Chouldechova, *The Effect of Differential Victim Crime Reporting on Predictive Policing Systems*, 2021 ACM Conference on Fairness, Accountability, and Transparency (2021), *available at* https://arxiv.org/abs/2102.00128.

386. *Feeding the Machine, supra* note 24, at 302.

387. Mark Harris, *How Peter Thiel's Secretive Data Company Pushed Into Policing,* Wired (Aug. 9, 2017), https://www.wired.com/story/how-peter-thiels-secretive-data-company-pushed-into-policing/.

388. *Id.*

389. *Id.*

390. *Id.*

391. "Similar to how your Facebook feed scours hundreds of friends' pages to produce a rolling digest of what it thinks are the most interesting posts, with this kind of algorithmic filtering, no human need ever look at the actual raw intelligence data. But as any Facebook user knows, such filters can produce results of wildly variable quality. And in police work, bad data can be dangerous." Harris *supra* note 387.

392. Richardson, *supra* note 95, at 202.

393. *Feeding the Machine, supra* note 24, at 300.

394. Caroline Haskins, *Academics Confirm Major Predictive Policing Algorithm Is Fundamentally Flawed*, Vice (Feb. 14, 2019, 12:57 PM), https://www.vice.com/en/article/xwbag4/academics-confirm-major-predictive-policing-algorithm-is-fundamentally-flawed [hereinafter Academics Confirm Major Predictive Policing Algorithm Is Fundamentally Flawed].

395. Task Force Interview with Michelle Fields at 83, Legal Aid Society-Supervising Attorney, Community Justice Unit, in N.Y.C. (Apr. 4, 2018).

396. Task Force Interview with Vincent Southerland at 309, Director of Race, Inequality, and the Law at NYU School of Law, in N.Y.C. (Apr. 18, 2018).

397. *Policing Predictive Policing supra,* note 15, at 1170.

398. Barrett, *supra* note 11, at 338.

399. *Id.* at 340.

400. Griffard, *supra* note 17, at 49.

401. Andrew Guthrie Ferguson, *The Allure of Big Data Policing*, PrawfsBlawg (May 25. 2017), https://prawfsblawg.blogs.com/prawfsblawg/2017/05/the-allure-of-big-data-policing.html.

402. Crawford & Schultz, *supra* note 204, at 105.

403. *The New Surveillance Discretion, supra* note 8, at 3.

404. Barrett, *supra* note 11, at 340.

405. *Stuck in a Pattern, supra* note 14.

406. Barrett, *supra* note 11, at 335.

407. Task Force Interview with Matt Cagle at 140-41, Technology and Civil Liberties Attorney, ACLU of Northern California, in S.F., Ca. (Feb. 11, 2019).

408. Griffard, *supra* note 17, at 52.

409. Barrett, *supra* note 11, at 340.

410. Barrett, *supra* note 11, at 344.

411. *Id.*

412. *PredPol's Stance on Privacy Civil Rights and Technology, supra* note 223.

413. Ellen Huet, *Serve and Protect, Predictive Policing Firm PredPol Promises to Map Crime Before It Happens*, Forbes (Mar. 2, 2015), https://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing/?sh=375938b24f9b.

414. *Academics Confirm Major Predictive Policing Algorithm Is Fundamentally Flawed, supra* note 390.

415. Task Force Interview with Philip Stark at 57, University of California, Berkeley (Jun. 27, 2019).

416. Stephine Seif et al., *Estimating ETAS: The Effects of Truncation, Missing Data, and Model Assumptions*, 22 JGR Solid Earth 499 (2016).

417. Task Force Interview with Nitin Kohli at 57, University of California, Berkeley (Jun. 27, 2019).

418. Alex Vitale, The End of Policing 28 (2017).

419. Ben Green, *The Just City: Machine Learning's Social and Political Foundations,* The Smart Enough City, (Mar. 29, 2019), https://smartenoughcity.mitpress.mit.edu/pub/vmjl8djz/release/1.

420. *Id.* at 416.

421. Matt Stroud, *Chicago's Predictive Policing Tool Just Failed a Major Test,* The Verge (Aug. 19, 2016, 10:28 AM), https://www.theverge.com/2016/8/19/12552384/chicago-heat-list-tool-failed-rand-test.

422. Harris, *supra* note 387.

423. Stop LAPD Spying, Before the Bullet Hits the Body – Dismantling Predictive Policing in L.A. 28, (May 8, 2018), https://stoplapdspying.org/wp-content/uploads/2018/05/Before-the-Bullet-Hits-the-Body-May-8-2018.pdf.

424. Harris, *supra* note 387.

425. Brayne, *supra* note 170.

426. Harris, *supra* note 387.

427. *Id.*

428. Alden, *supra* note 328.

429. Stanley, *supra* note 114.

430. *Policing Predictive Policing, supra* note 15, at 1170

431. Barrett, *supra* note 11, at 344.

432. *Feeding the Machine, supra* note 24, at 293.

433. *Id.*

434. Task Force Interview with Rebecca Wexler at 25, Professor at University of California, Berkeley School of Law and the Faculty Co-Director of the Berkeley Center for Law and Technology, in N.Y.C (Apr. 18, 2018).

435. Wexler, *supra* note 28, at 24.

436. *The Undue Influence of Surveillance Technology Companies on Policing, supra* note 48, at 120.

437. *Id.*

438. Gillmor, *supra* note 134.

439. *Id.*

440. *Id.*

441. Task Force Interview with Professor Elizabeth Joh at 210, U.C. Davis School of Law, in Washington D.C. (May 5, 2019). (Dr. Joh appeared via video teleconference).

442. Task Force Interview with Rashida Richardson at 194, Director of Police Research, AI Now Inst. at N.Y.U., in N.Y.C. (Apr. 18, 2018).

443. *The Undue Influence of Surveillance Technology Companies on Policing, supra* note 48, at 120.

444. Harris, *supra* note 387.

445. *The Undue Influence of Surveillance Technology Companies on Policing, supra* note 48, at 120.

446. "Automation is intensifying the privatization of the justice system. In recent years, private prisons have been found to undermaintain safety and security and private police have been found to operate with minimal training and oversight. The emerging criminal justice technologies discussed in this Article are also, for the most part, privately owned. 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, refusing to comply even with those subpoenas that seek information under a protective order and under seal." Wexler, *supra* note 28, at 1349.

447. *The Undue Influence of Surveillance Technology Companies on Policing, supra* note 48, at 125.

448. Josmar Trujillo & Alex S. Vitale, Brooklyn College, Gang Takedowns in the De Blasio Era – The Dangers of Precision Policing (2019), https://static1.squarespace.com/static/5de981188ae1bf14a94410f5/t/5df14904887d561d6cc9455e/1576093963895/2019+New+York+City+Gang+Policing+Report+-+FINAL%29.pdf.

449. Task Force Interview with Chad Marlow at 175, Advocacy and Policy Counsel, ACLU, in N.Y.C (Apr. 18, 2018).

450. Anderson, *supra* note 31, at 575.

451. *Id.* at 576.

452. *Gang Injunctions and Data, supra* note 36.

453. Muniz, *supra* note 33.

454. *Tracked and Trapped supra* note 35.

455. Salvador Hernandez, *A Database of Gang Members in California Included 42 Babies*, Buzzfeed News, Aug. 11, 2019, https://www.buzzfeednews.com/article/salvadorhernandez/database-of-gang-members-included-42-babies#.dnnryWodn

456. *Id.* at 5.

457. *N.Y. Gang Databases Expanded by 70% Under Mayor Bill De Blasio, supra* note 37.

458. *ACLU Demands Records on Boston's "Gang Database" Used in Deportations, supra* note 38.

459. Dooling, *supra* note 39.

460. Tracked and Targeted, *supra* note 40, at 5.

461. Anderson, *supra* note 31, at 582.

462. Jake Offenhartz, *The NYPD's Expanding Gang Database Is Latest Form of Stop & Frisk, Advocates Say,* The Gothamist (June 13, 2013, 3:00 PM), https://gothamist.com/news/the-nypds-expanding-gang-database-is-latest-form-of-stop-frisk-advocates-say.

463. Anderson, *supra* note 31, at 582.

464. *Private Eyes, They're Watching You: Law Enforcement's Monitoring of Social Media, supra* note 43.

465. Heffernan, *supra* note 339.

466. *Id.*

467. Howell, *supra* note 192, at 25.

468. Popper, *supra* note 330.

469. *Id.*

470. Trujillo & Vitale, *supra* note 448.

471. *Id.*

472. *Id.*

473. Howell, *supra* note 192, at 25.

474. *Id.*

475. Spec. Order S10-05, *supra* note 261.

476. Felton, s*upra* note 44.

477. Mick Dumke, *Chicago's Gang Database Is Full of Errors — And Records We Have Prove It,* ProPublica (Apr. 19, 2018, 4:00 AM), https://www.propublica.org/article/politic-il-insider-chicago-gang-database.

478. Hunt, *supra* note 280.

479. Felton, s*upra* note 44.

480. *N.Y. Gang Databases Expanded by 70% Under Mayor Bill De Blasio, supra* note 37.

481. *Tracked and Trapped, supra* note 35.

482. Policing in Chi. Research Grp., Univ. of Ill. at Chi., Accountability After Abolition (May 2019), *available at* http://erasethedatabase.com/2019/05/14/accountability-after-abolition/.

483. Olmstead v. United States, 277 U.S. 438, 478 (1928).

484. "The substance of all the definitions" of probable cause "is a reasonable ground for belief of guilt." McCarthy v. De Armit, 99 Pa. 63, 69 (1881) (quoted with approval in United States v. Carroll, 267 U.S. 132, 161 (1925). And this "means less than evidence which would justify condemnation" or conviction, as Marshall, C.J., said for the Court more than a century ago in *Locke v. United States*, 11 U.S. 339, 348 (1813). Since Marshall's time, at any rate, it has come to mean more than bare suspicion: Probable cause exists where "the facts and circumstances within their (the officers') knowledge and of which they had reasonably trustworthy information (are) sufficient in themselves to warrant a man of reasonable caution in the belief that' an offense has been or is being committed." Carroll, 267 U.S. at 162. *See* Brinegar v. United States, 338 U.S. 160.

485. The Fourth Amendment permits brief investigative stops — such as the traffic stop in this case — when a law enforcement officer has "a particularized and objective basis for suspecting the particular person stopped of criminal activity." United States v. Cortez, 449 U.S. 411, 417– 418 (1981); *see also* Terry v. Ohio, 392 U. S. 1, 21–22 (1968). The "reasonable suspicion" necessary to justify such a stop "is dependent upon both the content

of information possessed by police and its degree of reliability." Alabama v. White, 496 U. S. 325, 330 (1990). The standard takes into account "the totality of the circumstances — the whole picture." Cortez, *supra,* at 417. Although a mere "hunch" does not create reasonable suspicion, Terry, *supra,* at 27, the level of suspicion the standard requires is "considerably less than proof of wrongdoing by a preponderance of the evidence," and "obviously less" than is necessary for probable cause. United States v. Sokolow, 490 U. S. 1, 7 (1989); Navarette v. California, 572 U.S. 393 (2014).

486. *The Undue Influence of Surveillance Technology Companies on Policing, supra* note 48, at 134.

487. "For a second level of inquiry, imagine the police officer uses networked databases owned by third parties to discover personal information about a suspect. This data might include credit information, financial records, credit card activity, employment, past addresses and telephone numbers, names and addresses of family members, neighbors' addresses and telephone numbers, business associates, make, model, and color of registered vehicles, social security numbers, dates of birth, bankruptcies, liens and judgments, and GPS locational data. While access to some of these data would usually require particular legal authorization, law enforcement can circumvent statutes restricting direct access by instead using 'fourth-party' commercial aggregators." *Id.* at 379.

488. *The New Surveillance Discretion, supra* note 8, at 28.

489. "If data are collected only about certain classes of people, then those people are more likely to become future targets of suspicion simply because of the initial selection bias. Thus, important questions remain about who collects, interprets, and chooses the big data to study. Worse, like other quantitative systems used for decision-making, big data-based predictive policing will appear to be objective and fair when it may in fact reflect subjective factors and structural inequalities. Just as we have credit ratings that allow lenders to predict future creditworthiness, police could develop "criminal ratings" to predict future criminal proclivity." *Predictive Policing and Reasonable Suspicion, supra* note *89, at 402.*

490. Unvalidated and unreliable forensic evidence is undermining criminal trials. In 2009, a National Academy of Sciences report identified a "notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods" and noted that numerous forensic disciplines lack known accuracy measures or error rates. Wexler, *supra* note 28, at 1421.

491. *Id.*

492. 373 U.S. 83 (1963).

493. Fields, *supra* note 395.

494. Unvalidated and unreliable forensic evidence is undermining criminal trials. In 2009, a National Academy of Sciences report identified a "notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods" and noted that numerous forensic disciplines lack known accuracy measures or error rates. Wexler, *supra* note 28, at 1421.

495. United States v. Ellis, No. 19-369 (W.D. Pa. Feb. 26, 2021), *available at* https://www.courtlistener.com/recap/gov.uscourts.pawd.262237/gov.uscourts.pawd.262237.138.0_1.pdf.

496. U.S. v Ellis, no. 19-369, (W.D. Pa. Feb. 26, 2021), *available at* https://www.courtlistener.com/recap/gov.uscourts.pawd.262237/gov.uscourts.pawd.262237.138.0_1.pdf

497. Pickett, 246 A.3d at 324.

498. *Racial Equity in Algorithmic Criminal Justice, supra* note 51, at 1088.

499. "Discriminatory intent" is a central term in the judicial interpretation of constitutional clauses requiring the equal treatment of persons without regard to their race, ethnicity, or religion. Aziz Z. Huq, *What Is Discriminatory Intent*, 103 Cornell L. Rev. 1211, 1212 (2018).

500. Vincent Southerland, *The Intersection of Race and Algorithmic Tools in the Criminal Legal System*, 80 Md. L. Rev. (forthcoming 2021), *available at* https://ssrn.com/abstract=3797102.

501. "(I)ncreasing evidence suggests that human prejudices have been baked into these tools because the machine-learning models are trained on biased police data. Far from avoiding racism, they may simply be better at hiding it. Many critics now view these tools as a form of techwashing, where a veneer of objectivity covers mechanisms that perpetuate inequities in society." Will D. Heaven, Predictive Policing Algorithms Are Racist. They Need to Be Dismantled, MIT Tech. Rev., (July 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/.

502. *Racial Equity in Algorithmic Criminal Justice supra* note 51, at 1076.

503. *The Undue Influence of Surveillance Technology Companies on Policing supra* note 48, at 132.

504. "The constitutional requirement that warrants must particularly describe the 'things to be seized' is to be accorded the most scrupulous exactitude when the 'things' are books, and the basis for their seizure is the ideas which they contain." [Footnote 16] See Marcus v. Search Warrant, 367 U. S. 717; A Quantity of Copies of Books v. Kansas, 378 U. S. 205. No less a standard could be faithful to First Amendment freedoms. 379 U.S. 476 (1965).

505. Daniels *supra* note 336.

506. Task Force Interview with Taylonn Murphy at 89, Community Organizer and Activist, in N.Y.C, (Apr. 18, 2018).

507. Stevie Degroff & Albert Fox Cahn, *New CCOPS on the Beat, Stop Surveillance Tech.* Oversight Project, 1, (Feb. 10, 2021), https://static1.squarespace.com/static/5c1bfc7eee175995a4ceb638/t/602430a5ef89df2ce6894ce1/1612984485653/New+CCOPS+On+The+Beat.pdf.

508. *Id.* at 2.

509. Task Force Interview with Malkia Cyril at 211, Executive Director and founder of the MediaJustice, in San Francisco.

510. Davide Castelvecchi, *Mathematicians Urge Colleagues to Boycott Police Work in Wake of Killings*, Nature, Jun. 19, 2020, https://www.nature.com/articles/d41586-020-01874-9

511. Leila Miller, *LAPD will end Controversial Program that Aimed to Predict where Crimes would Occur*, L.A. Times, (Apr 21, 2020), https://www.latimes.com/california/story/2020-04-21/lapd-ends-predictive-policing-program.

512. George Joseph, *Majority of Manhattan DA Candidates Pledge to Sever Ties with Palantir*, The Gothamist, (Feb 18, 2021), https://gothamist.com/news/majority-manhattan-da-candidates-pledge-sever-ties-palantir

513. Degroff & Cahn *supra* note 507, at 2.

514. *Id.*

515. *Id.*

516. *Racial Equity in Algorithmic Criminal Justice supra* note 51, at 1065.

517. Task Force Telephone Interview with Cathy O'Neil at 11, O'Neil Risk Consulting & Algorithmic Auditing (Oct. 17, 2018).

518. Cyril *supra* note 509.

519. Murphy *supra* note 506.

520. Task Force Interview with Scott Levy at 106, Special Counsel, Criminal Defense Practice, The Bronx Defenders. in N.Y.C., (Apr. 18, 2018).

521. Ruha Benjamin, *Race After Technology*, 5-6 (2019).

522. Southerland, *supra* note 396.

523. The Task Force defines "data-driven policing" as including, but not limited to, the surveillance technologies, tools, and methods employed by law enforcement to visualize crime; target "at-risk" individuals and groups; map physical locations; track digital communications; and collect data on individuals as well the communities they patrol. "Data-driven policing" also encompasses place-based predictive models that rely on historical crime data, geographic data, and demographic data; person-based predictive models that rely on personal data and social network analysis; and any databases, lists, and systems that subject individuals to increased police surveillance and monitoring.

524. Brief of Amici Curiae Public Justice Center, American Civil Liberties Union of Maryland, and Washington Lawyers' Committee for Civil Rights and Urban Affairs, *Sizer v. Maryland* (Md. Ct. App.) 9 (2017) (*quoting* Kelly Koss, *Leveraging Predictive Policing Algorithms to Restore Fourth Amendment Protections in High-Crime Areas in a Post-Wardlow World*, 90 Chi-Kent L. Review 301, 312 (2015).

525. Wexler, *supra* note 28, at 1368.

526. Wexler, *supra* note 28, at 1429.

527. When new surveillance technologies are kept secret because of non-disclosure agreements, they cannot be challenged by criminal defendants and these challenges cannot be decided by judges, regardless of the merits of the defendants' claims. The use of a new surveillance technology may or may not be considered a Fourth Amendment search, but a private company's insistence on secrecy removes the legal issue from judicial review.

528. Laura M. Moy, *A Taxonomy of Police Technology's Racial Inequity Problems*, 2021 U. Ill. L. Rev. 139 (2021).

529 *Tracked and Trapped, supra* note 35.

530. Mariam Arain et al., *Maturation of the Adolescent Brain*, 9 Neuropsychiatric Disease and Treatment 449, 453 (2013).

531. Text in italics indicates the title of the witness at the time of their meeting with the Task Force.

532. *See generally*, Ari Chivukula & Tyler Takemoto, Samuelson Law, Tech. & Pub. Pol'y Clinic, Local Surveillance Oversight Ordinances (Feb. 2021), https://www.law.berkeley.edu/wp-content/uploads/2021/02/Local-Surveillance-Ordinances-White-Paper.pdf; *see also* Community Control Over Police Surveillance, ACLU, https://www.aclu.org/issues/privacy-technology/surveillance-technologies/community-control-over-police-surveillance#map (last visited Mar. 12, 2021).

533. S.F., Cal., Admin. Code ch. 19B (2019).

534. S.F. Bay Area Rapid Transit District, Cal., Code ch. 17, art. V (2018)

535. Oakland, Cal., Mun. Code ch. 9.64 (2021).

536. Berkeley, Cal., Mun. Code ch. 2.99 (2018).

537. Davis, Cal., Mun. Code art. 26.07 (2018).

538. Palo Alto, Cal., Mun. Code §§ 2.30.620-2.30-690 (2018).

539. San Diego, Cal., Ord. Nos. O-2021-67, O-2021-69 (*passed on first reading* Nov. 10, 2020). Note: The City Council must give both ordinances final approval before they take effect.

540. Santa Clara Cnty., Cal., Mun. Code tit. A, div. 40 (2016).

541. New Orleans, La., Mun. Code ch. 147 (2020).

542. Cambridge, Mass., Mun. Code ch. 2.128 (2018).

543. Lawrence, Mass., Mun. Code ch 9.25 (2018).

544. Somerville, Mass., Mun. Code pt. II, ch. 10, art. III (2019).

545. Grand Rapids, Mich., Admin. Pol'y No. 15-03 (2015).

546. N.Y.C., N.Y., Admin. Code § 14-188 (2020).

547. Yellow Springs, Ohio, Mun. Code ch. 607 (2018).

548. Pittsburgh, Penn., Admin. Code § 116.15 (2020).

549. Nashville, Tenn., Metro. Code § 13.08.080 (2017).

550. Seattle, Wash., Mun. Code ch 14.18 (2017).

551. Madison, Wis., Mun. Code §§ 23.63-23.64 (2020).

552. Kristi Sturgell, *Santa Cruz Becomes the First U.S. City to Ban Predictive Policing*, L.A. Times (June 26, 2020), https://www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing.

553. Keith Burbank, *City First in Nation to Ban Predictive Policing, Biometric Surveillance Tech*, SFGate (Jan. 14, 2021), https://www.sfgate.com/news/bayarea/article/City-First-In-Nation-To-Ban-Predictive-Policing-15872642.php.

554. Michael Stein, *New Orleans City Council Bans Facial Recognition, Predictive Policing and Other Surveillance Tech*, The Lens (Dec. 18, 2020), https://thelensnola.org/2020/12/18/new-orleans-city-council-approves-ban-on-facial-recognition-predictive-policing-and-other-surveillance-tech/.

555. Mark Puente, *LAPD Pioneered Predicting Crime with Data. Many Police Don't Think It Works*, L.A. Times (July 3, 2019), https://www.latimes.com/local/lanow/la-me-lapd-precision-policing-data-20190703-story.html.

556. Ryan Deto, *Pittsburgh City Council Introduces Police Facial Recognition, Predictive Policing Ban*, Pittsburgh City Paper (Aug. 25, 2020), https://www.pghcitypaper.com/pittsburgh/pittsburgh-city-council-introduces-police-facial-recognition-predictive-policing-ban/Content?oid=17879052.

557. Ian Bauer, *Milpitas, Calif., Police Department Nixes Predictive Policing Contract*, Gov't Tech. (July 14, 2016), https://www.govtech.com/public-safety/Milpitas-Calif-Police-Department-Nixes-Predictive-Policing-Contract.html.

558. Puente, *supra* note 555.

559. *Id.*

560. *Id.*

Reckless Automation in Policing

__ Berkeley Tech. L. J. __ (2022)

Elizabeth E. Joh[1]

Introduction

Automated decision-making plays an increasingly larger role in policing.[2] Traditional methods of police investigation have been augmented by tools like facial recognition, predictive analytics, license plate readers, and robotics.[3] These tools allow the police to sift through large amounts of information at a scale and speed not practicable with human skills alone. This reliance on artificial intelligence, however, has prompted numerous questions about how to balance criminal investigation needs with concerns about fairness, bias, transparency, and accountability. These concerns aren't unique to policing. You can find similar calls for "algorithmic accountability" in healthcare, banking, credit scoring, public benefits, and employment.[4]

How should we evaluate the growth of automation in policing? There is no shortage of answers, but this essay starts with a simple observation: by focusing on automation's

---

[1] Professor Law of Law, U.C. Davis School of Law. Thanks to the members of the Berkeley Technology and Law Journal for organizing the 2021 symposium, "Technology Law as a Vehicle for Anti-Racism."

[2] A recent report on predictive policing summarized the current use of technology in policing this way: "As the cost of collecting, storing, and analyzing data falls to nearly zero, we should expect a proliferation of data analysis tools and algorithmic mediation between the citizen and the state." See Ethics + Emerging Sciences Group, Artificial Intelligence + Predictive Policing: An Ethical Analysis 23 (2020).

[3] Cf. Andrew G. Ferguson, Policing Predictive Policing, 94 Wash. U. L. Rev. 1109, 1114 (2017) (noting that "the first generation of predictive policing technologies represents only the beginning of a fundamental transformation of how law enforcement prevents crime"). A partnership between the Electronic Frontier Foundation and the University of Nevada, Reno Reynolds School of Journalism has produced the Atlas of Surveillance an ongoing project that maps the use of law enforcement technologies like predictive policing, face recognition, and license plate readers. See EFF, Atlas of Surveillance (2019), at https://atlasofsurveillance.org/.

[4] There is already a large literature on algorithmic accountability and transparency, with many different approaches. Some notable examples include: Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249 (2008); Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. Rev. 1, 20 (2014); Tal Zarsky, Transparent Predictions, 2013 U. Ill L. Rev. 1503 (2013); Devin R. Desai & Joshua A. Kroll, Trust But Verify: A Guide to Algorithms and the Law, 31 Harv. J.L. & Tech. 1 (2017); Joshua A. Kroll et al., Accountable Algorithms, 165 U. Pa. L. Rev. 633 (2017).

harms to persons first. American policing is rife with reckless automation. The highly decentralized system of policing in the United States, with its more than 18,000 agencies, permits and encourages experimentation with new technologies.[5] Innovation in policing can, of course, be positive. Crime control and public safety are complex and evolving social problems, and over time, the police change their tactics and tools to address them.

But new technologies that rely on artificial intelligence and the vast amounts of digital information now available have introduced new problems.[6] Police departments have bought, licensed, adopted, and experimented with technologies that impact communities through increased but invisible surveillance, and with mistakes that impose real-life consequences in police-civilian interactions. And these technological experiments are often deployed in places or against communities that have often already been overpoliced. Those who are disproportionately and frequently affected by these experiments are black, brown, low-income, and without significant political power. We should identify this development as reckless automation in policing.[7]

Reckless automation has tangible consequences: its mistakes lead to street stops, arrests, and traffic stops of individuals. Communities also experience the psychological costs of pervasive (and barely visible) automated surveillance. Sometimes these policing experiments are conducted without the knowledge of the communities involved.

If we accept the premise of reckless automation, the conversation about accountability, artificial intelligence, and policing might benefit from a seemingly unrelated policy framework: that of experimentation on human subjects. The comparison may seem far-fetched. Yet even the police may think of their new technologies in the same way that the medical community approaches experimentation. The Los Angeles Police Department, for instance, referred to one of their uses of artificial intelligence as "dosage."[8] Borrowing

[5] Bureau of Justice Statistics, National Sources of Law Enforcement Employment Data 1 (2016) ("Law enforcement in the United States is made up of about 18,000 federal, state, county, and local agencies.").
[6] Cf. David G. Robinson and Miranda Bogen, Upturn, Automation and the Quantified Society 9 (2017) ("Governments around the world increasingly use automation to make important decisions about people's lives, often without broad consultation or careful assessment of new systems' impact.").
[7] Reckless here is used in the criminal law sense: the conscious disregard of a substantial and unjustifiable risk. See Model Penal Code 2.02(c).
[8] In its 2019 review of the PredPol predictive policing software used by the L.A.P.D., the Los Angeles Inspector General noted that "the amount of time an officer spends in a PredPol hotspot is referred to as

from that framework does not imply that reckless automation in policing is the literal equivalent of medical or psychological experiments on human subjects. Nor does such a comparison imply that the technical aspects of institutional review boards should apply directly to new policing strategies.[9] But turning to a bioethical framework has value because it draws attention to the *subjects*—the communities affected--of policing. To the extent that the ethical considerations applied in human subjects research provide useful insights to apply to the many changes in policing, they open a new conversation. What if we think of new forms of automated decision-making in policing as experiments on communities that might impose harms with life-altering decisions?[10]

Automation and Accountability

We all know about the influence of artificial intelligence and automation in our lives, from the most mundane experiences, like picking favorite songs or movies, to more important decisions like who should receive job interviews or who should receive government benefits.[11] That automation is also a part of policing. License plate readers today use algorithms to quickly identify individual plates. Facial recognition technology can quickly identify faces in fixed databases or in real-time scans. Software identifies high risk persons and places.

All of this automation falls under the umbrella of "artificial intelligence": the use of machines to assume cognitive tasks usually performed by people.[12] That is a very broad definition, and rightly so: artificial intelligence can involve everything from the

dosage, which an be measured in minutes or hours." Ultimately, the review concluded that the effectiveness of PredPol based "dosage" was inconclusive. Office of the Inspector General, Review of Selected Los Angeles Police Department Data-Driven Policing Strategies 25, 29 (2019).

[9] And in fact, federal regulations on human subject research explicitly excludes data collected for "criminal investigative purposes." 45 C.F.R. 46.101 (l)(4)

[10] Cf. David G. Robinson and Miranda Bogen, Upturn, Automation and the Quantified Society 11 (2017) (describing "life-altering" effects of automation)

[11] Artificial intelligence is an umbrella term for cognitive tasks that have been assumed by machines, and one form of artificial intelligence, machine learning, involves programming computers to detect patterns in the data provided. *See, e.g.,* Calo, supra note xx, at 404.

[12] *See,* Upturn, supra note xx, at 11.

application of straightforward algorithms[13] to more complicated examples of machine learning that learns to identify patterns in data.[14] Some artificial intelligence simply provides more information for human decisionmakers (risk assessments in finance[15]), while other forms perform the analysis and the action (hiring and employment decisionmaking).[16]

But automation also poses questions about bias, secrecy, unaccountability, and mistakes that are hard to spot when the decision originates from a machine and not a person.[17] While the United States lacks a comprehensive data protection regime, many proposals regulatory proposals have been proposed and some have become law. These proposals can apply to automated decisionmaking generally, or to more specific subject matters like policing and criminal justice.

We can summarize some of the predominant approaches to ethics and accountability in artificial intelligence.

First, because machine learning involves the identification of patterns from enormous data sets, the constitution of that training data can be a problem.[18] If a dataset of faces

---

[13] Algorithms are "simply a sequence of steps used to accomplish some task." Upturn, supra note xx, at 13.
[14] While artificial intelligence has been researched since the 1940s, the importance and pervasiveness of AI today is a result of 1) the availability of huge amounts of data; 2) improvements in machine learning; and 3) improvements in computing power. *See* Congressional Research Service, Artificial Intelligence and National Security 2 (2020).
[15] See, e.g., OECD, Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities Challenges and Implications for Policy Makers 29 (2021), at https://www.oecd.org/finance/financial-markets/Artificial-intelligence-machine-learning-big-data-in-finance.pdf ("AI-based models and big data are increasingly being used by banks and fintech lenders to assess the creditworthiness of prospective borrowers and make underwriting decisions.").
[16] *See, e.g.,* Spencer Soper, *Fired by Bot at Amazon: "It's You Against the Machine,"* Bloomberg, June 28, 2021, at https://www.bloomberg.com/news/features/2021-06-28/fired-by-bot-amazon-turns-to-machine-managers-and-workers-are-losing-out ("At Amazon, machines are often the boss-hiring, rating and firing millions of people with little or no human oversight.").
[17] Cf. Ryan Calo, Artificial Intelligence Policy: A Primer and Roadmap, 51 U.C. Davis, L. Rev. 399, 407 (2017) (noting "AI presents unique and important ethical questions").
[18] Solon Barocas and Andrew D. Selbst, Big Data's Disparate Impact, 104 Cal. L. Rev. 671, 680 (2016) (defining training data as "quite literally the data that train the model to behave in a certain way"); Nicol Turner Lee, et al., Algorithmic Bias Detection and Mitigation: Best practices and policies to reduce consumer harms, Brookings, May 22, 2019, at https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/ ("If the data used to train the algorithm are more representative of some groups of people than others, the predictions from the model may also be systematically worse for unrepresented or under-represented groups."); Xavier Ferrer et al., Bias and Discrimination in AI: a cross-disciplinary perspective, __ Computers & Soc'y __ (2020), at https://arxiv.org/pdf/2008.07309.pdf (noting "algorithms learn to make decisions or

has many more white people then non-white people, then a facial recognition program instructed to identify faces can misidentify those who are not white at much higher rates than whites.[19] One study of facial recognition algorithms found that that while white men were correctly identified nearly all the time, black women were incorrectly identified up to a third of the time.[20] Put simply, biased data will lead to biased results.[21] Thos concern has prompted both calls for better training data and for bans or pauses on the use of facial recognition technology until these problems have been addressed.[22]

Relatedly, there can be bias in the algorithms themselves. People create algorithms, and their assumptions about the appropriate design and execution of the algorithm may create further biases. A recruitment algorithm, for instance, that uses men as the model for professional "fit" will disadvantage female applicants.[23] Of course, bias is not a new idea. But in this context, bias--whether in training data or in the instructions themselves--can magnify inequalities by reproducing these effects on a very large scale. Some scholars have proposed as a solution public access both to source codes and data sets.[24]

Another proposal in artificial intelligence policy is the call for explainability. With some complex uses of artificial intelligence, programmers may not be able to explain exactly why or how particular outcomes have been achieved. This black box problem means that a person may not be able to know why a particular prediction or decision was made.[25]

---

predictions based on datasets that often contain past decisions. If a dataset is used for training purposes reflects existing prejudices, algorithms will very likely learn to make the same biased decisions.").

[19] Natasha Singer, Amazon is Pushing Facial Technology that a Study Says Could Be Biased, N.Y. Times, Jan. 24, 2019, at https://www.nytimes.com/2019/01/24/technology/amazon-facial-technology-study.html.

[20] Joy Buolamwini & Timnit Gebru, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification, 81 Proceedings of Machine Learning Research 1 (2018), at http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

[21] See, e.g. Citron & Pasquale, supra note xx, at 4 ("Scoring systems mine datasets containing inaccurate and biased information provided by people.").

[22] Turner Lee, supra note xx.

[23] See id.

[24] See e.g., Danielle Keats Citron, Technological Due Process, 85 Wash. U. L. Rev. 1249, 1308 (2008); Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions, 89 Wash. L. REv. 1, 20, 26 (2014) ("the logics of predictive scoring systems should be open to public inspection").

[25] See, e.g., Jessica Newman, Explainability won't save AI, Brookings, May 19, 2021, at https://www.brookings.edu/techstream/explainability-wont-save-ai/ ("Much of artificial intelligence . . . is plagued by the "black box problem." While we may know the inputs and outputs of a model, in many cases we do not know what happens in between.").

Thus, there have been calls for a "right to explanation" when, for example, a person receives an adverse employment decision through automation.[26] And although calling for the explanation of why an automated decisionmaking process led to a person's rejection for a loan, a job, or for benefits has appeal, there is not yet widespread consensus on what form that explainability should take.[27] The right to an explanation can be combined with other tools from a legal framework to provide individuals with "technological due process" rights in automated decisionmaking.[28]

Another critique of secrecy in automated decisionmaking arises because these tools are often developed within the private sector.[29] Government entities, including law enforcement agencies, typically do not design or create these systems.[30] Instead, public agencies usually stand in a customer-vendor relationship with private companies and then adopt the tools of algorithmic decisionmaking as a matter of purchase, lease, or contract.[31] These relationships complicate accountability considerably. If a person receives an adverse decision for government benefits because of a prediction tool developed privately, the agency may be unable to provide an explanation for the reasoning because the vendor invokes its propriety interests and refuses to provide information. Criminal defendants have encountered problems, for example, in trying to access the source code for the privately developed "probabilistic typing" software that analyzes DNA

[26] The EU's General Data Protection Regulation (GDPR), for instance, provides for a "right to be informed" about algorithmic decisionmaking. For a comprehensive analysis of the GDPR right to explanation, see Margot E. Kaminski, The Right to Explanation, Explained, 34 Berkeley Tech. L. J. 190, 209-217 (2019).

[27] Upturn, supra note xx, at 15.

[28] Citron, supra note xx, at 1306 (arguing for a due process framework in automating decisionmaking, including the right of explanation).

[29] *See, e.g.,* Rebecca Wexler, Life, Liberty, and trade secrets: Intellectual Property in the Criminal Justice System, 70 Stan. L. Rev. 1343 (2018)(arguing that trade secrets regarding criminal justice technologies should not be privileged in criminal proceedings).

[30] Hanna Bloch-Webha, Visible Policing: Technology, Transparency, and Democratic Control, 109 Cal. L. Rev. __ (2021)(noting "new technologies of surveillance, often procured from other otherwise reliant on the private sector, tend to operate in opaque and unaccountable ways").

[31] See, e.g., Elizabeth E. Joh, The Undue Influence of Surveillance Technology Companies on Policing, 92 N.Y.U. L. Rev. Online 101 (2017)(discussing influence of private companies producing surveillance hardware and software on democratic policing).

samples that are usually too difficult for traditional forensics labs to assess but has nevertheless identified them as a suspect.[32]

These approaches to algorithmic accountability each identify important problems in the uses of automated decisionmaking. Each proposes solutions that can be implemented in new regulations and agency decisionmaking. And all have influenced efforts to regulate automated decisionmaking around the country. In policing, concerns about secrecy, for instance, have led some local governments to impose notice and reporting requirements on news uses of surveillance technologies by their police departments.[33] But all of these share a similar premise: that algorithmic accountability address a *technological process* that requires new forms of regulation in order to be implemented fairly. These are solutions for fixing machines. What if we started somewhere else?

A Different Framing: Experimentation

Biomedical and behavioral research involving human subjects today is generally subject to institutional review boards focused on the potential ethical consequences of that work. Federally funded research must abide by federal regulations regarding human subjects, including informed consent procedures.[34] Human subjects refer to any "living individual about whom an investigator" then obtains "information or biospecimens through intervention or interaction with the individual and uses, studies, or analyzes the information," or does the same for "identifiable private information."[35] Additionally,

---

[32] See, e.g., Lauren Kirchner, Where Traditional DNA Testing Fails, Algorithms Take Over, ProPublica, Nov. 4, 2016, at https://www.propublica.org/article/where-traditional-dna-testing-fails-algorithms-take-over ("Defendants' requests to get access to TrueAllele's source code have consistently been denied."); but also Lauren Kirchner, Powerful DNA Software Used in Hundreds of Criminal Cases Faces New Scrutiny, The Markup, Mar. 9, 2021, at https://themarkup.org/news/2021/03/09/powerful-dna-software-used-in-hundreds-of-criminal-cases-faces-new-scrutiny.

[33] *See, e.g.*, Ari Chivukula & Tyler Takemoto (Samuelson Clinic), Local Surveillance Oversight Ordinances (2021), at https://www.law.berkeley.edu/wp-content/uploads/2021/02/Local-Surveillance-Ordinances-White-Paper.pdf (counting sixteen American jurisdictions that have "passed local surveillance technology oversight ordinances meant to bring more transparency and democratic control to local government use of surveillance technology").

[34] The Federal Policy for the Protection of Human Subjects, first published in 1991, is also referred to as the "Common Rule" and can be found in 45 C.F.R. 46 of the Department of Health and Human Services (HHS) regulations. The Common Rule was updated in 2018. See, e.g., Jerry Menikoff et al., The Common Rule, Updated, 376 New England J. Med. 613 (2017).

[35] C.F.R. 46.102 (e)(1).

research is defined as "a systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge."

Heavily influential to the system for protecting human research subjects today is a report written in 1978 by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.[36] Also known as the Belmont Report, the Commission's work was prompted by notorious abuses of human research subjects, including the 1972 reporting on the infamous Tuskegee Study.[37] In the infamous experiment, the U.S. Public Health Service offered to treat 600 African American men "for bad blood" in exchange for meals, medical exams, and burial insurance.[38] They were not informed that the actual purpose of the study was to examine the effects of untreated syphilis, and were denied access to a cure. In writing the Belmont Report, the Commission took the view that "risk-laden, albeit promising research" might not be justified "merely on the strength of its potential social benefits."[39]

The hallmarks of the Belmont Report are its three fundamental principles: respect for persons, beneficence, and justice. A respect for persons includes the assumption that "individuals should be treated as autonomous agents" whose "considered opinions and choices" are entitled to respect.[40] Beneficence requires efforts to secure the "well-being" of research subjects, including maximizing possible benefits and minimizing possible to harms to them.[41] The third principle the Report emphasizes is justice: that "an injustice occurs when some benefit is entitled without good reason or when some burden is imposed unduly."[42]

---

[36] Eli Y. Adashi et al., The Belmont Report at 40: Reckoning With Time, 108 Am. J. Public Health 1345 (2018).

[37] Also influential before the Belmont report were the Nuremberg code and the Declaration of Helsinki: each a framework for ethical considerations in research on human subjects. See, e.g., Kaille Kodama Muscente, Ethics and the IRB: the History of the Belmont Report, Aug. 3, 2020, at https://www.tc.columbia.edu/institutional-review-board/irb-blog/the-history-of-the-belmont-report/

[38] Centers for Disease Control Control and Prevention, The Tuskegee Timeline (2021), at https://www.cdc.gov/tuskegee/timeline.htm.

[39] See Adashi, supra note xx, at 1345.

[40] The National Commission for the Protection of Human Subjets of Biomedical and Behavioral Research, The Belmont Report ("Belmont Report") 4 (1979).

[41] Belmont Report, supra note xx, at 5.

[42] Belmont Report at 6.

For human subjects research, theses three principles translate into practical steps: the use of informed consent by research subjects, a risk/benefit assessment about whether to perform the research, and the careful selection of subjects. As to this final concern, the Belmont Report raises this note of caution:

> Injustice may appear in the selection of subjects . . . Thus injustice arises from social, racial, sexual and cultural biases institutionalized in society. Thus, ... unjust social patterns may nevertheless appear in the overall distribution of the burdens and benefits of research. . . .[43]

These three considerations—respect for persons, beneficence, and justice—are useful rubrics for thinking about the ethics of technology in policing. The technologies on which the police increasingly rely all promise advances in how investigations are conducted. But their costs, whether through mistaken arrests and stops or pervasive surveillance have yet to find a meaningful ethical framework.

Automation, Bioethics, Policing

How can such a very different conversation on law and policy regarding human subject research inform the one in police automation? The algorithmic accountability movement has offered many proposals to address the problems of automation. But these conversations focus first on the technology: how to modify, reform, and monitor both the data and design of automated decisionmaking. All of this remains important. But to pose the question as the beginning of this essay: what if we think of new forms of automated decisionmaking in policing as experiments on communities that might impose harms with life altering decisions?[44]

First, we know that surveillance of all kinds is distributed unevenly in society. Receipt of public benefits can often require subjection to drug tests, fingerprinting, verification

[43] Belmont Report at 9.

[44] Cf. Kate Crawford & Ryan Calo, There is a blind spot in AI research, Nature, Oct. 13, 2016, at https://www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805 (noting "there are no agreed methods to assess the sustained effects of such applications on human populations").

requirements, and privacy-intrusive questions.[45] Non-white, low income communities are more often subjected to heavy-handed police presence and surveillance than their wealthier and whiter counterparts.[46] Online, low-income Americans face disadvantages because they usually buy less expensive digital devices with fewer privacy protections and possess fewer digital skills to keep their information private.[47] Low wage work is particularly subject to tracking about movements, productivity, drug tests, and other forms of surveillance.[48]

In addition, non-white, low-income communities are not just subjected to more surveillance, but also more *combinations* of surveillance than other groups.[49] The potential harms from living with pervasive and inescapable surveillance are quite real.[50] People in low income communities can find it difficult to protect their privacy and to correct mistakes that lead to adverse decisions in housing, credit, and policing. This means some communities live with less autonomy and freedom from surveillance than other groups do.

In the case of policing, we might consider the decision to "test" out a new automated decisionmaking on a community as a form of experimentation with impacts that might benefit from the ethical concerns in human subjects research. When a law enforcement agency decides to pilot or adopt automated decisionmaking today, it might do so without informing or receiving consent or input from the community policed; without explicit consideration of whether the experiment maximizes benefits while minimizing harms; or

---

[45] Khiara M. Bridges, Privacy Rights and Public Families, 34 Harv. J.L. & Gender 113, 114-116 (2011)(discussing Medicaid); Virginia Eubanks, Want to Predict the Future of Surveillance? Ask Poor Communities, The American Prospect, Jan. 15, 2014, at https://prospect.org/power/want-predict-future-surveillance-ask-poor-communities./ ("Poor and working class Americans already live in the surveillance future.").

[46] Lori Beth Way & Ryan Patten, Hunting for "Dirtbags" 3 (2013) ("through discretionary proactive policing (done in their 'free time'), law enforcement officers monitor the lower classes to a greater degree than the middle and upper classes. Such police behavior feeds the cycle of depositing the poor into the criminal justice system and ensuring they remain under criminal justice scrutiny").

[47] See Mary Madden et al., Privacy, Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans, 95 Wash. L. Rev. 53, 62 (2017) (noting that the poor are increasingly online, but face privacy and security disadvantages).

[48] Madden, supra note xx, at 60.

[49] Madden, et al., at 63 ("It is important to recognize that for the poor, overt and covert surveillance systems interact with one another.").

[50] Madden, at 61.

without considering whether a program unduly burdens that community. We would not condone an experiment on a community with potential psychological and physical impacts without ethical approvals. But none of the ethical principles fundamental to human subjects research are usually considered for new policing technologies.

The absence of such ethical considerations are striking when we know that police departments do in fact experiment with and sometimes have abandoned automated decisionmaking programs that significantly impact the communities affected. Consider two recent examples.

In 2012, the Chicago Police Department began to use risk models, popularly described as its "Heat Lists,"[51] to identify those who were likely to become victims or perpetrators of gun violence within the next eighteen months.[52] A research team from the Illinois Institute of Technology created the risk models and calculated the "scores" for individuals--everyone who had been arrested four years before the calculations were made.[53] The higher the risk score, the higher chance that a person would become a "Party to Violence," either as a victim or perpetrator of gun violence.[54] These scores were available to all Chicago Police Department personnel, as well as on its crime mapping software.[55] By 2018, there were nearly 400,000 people with individualized risk scores under the program. The majority of black men in Chicago between the ages of 20 and 29 had a risk score under the program.[56]

It its 2020 report, the City of Chicago's Inspector General found the department's predictive program filled with "concerns."[57] The individualized scores and risk tiers used in the program were found to be "unreliable." In addition, the scores, premised on arrests

---

[51] John S. Hollywood (RAND), CPD's "Heat List" and the Dilemma of Predictive Policing, Sept. 21, 2016, at https://www.rand.org/blog/2016/09/cpds-heat-list-and-the-dilemma-of-predictive-policing.html (noting Strategic Subjects List is "known colloquially as the 'heat list'").

[52] City of Chicago Office of Inspector General, Advisory Concerning the Chicago Police Department's Predictive Risk Models ("OIG Report") 1 (2020).

[53] Id. at 2.

[54] Id.

[55] Id.

[56] Yana Kunichoff & Patrick Sier, The Contradictions of Chicago Police's Secretive List, Chicago Magazine, August 21, 2017, at https://www.chicagomag.com/city-life/august-2017/chicago-police-strategic-subject-list/.

[57] OIG Report, supra note xx, at 4.

that did not necessarily lead to conviction, may have been the basis of police interventions that "effectively punished individuals for criminal acts for which they had not been convicted."[58] Having a high risk score may have led to some people receiving harsher charging decisions in subsequent arrests, even if they had never been convicted for the prior arrest.[59] The Department formally decommissioned the program in November 2019.[60]

The risk assessment program used in Chicago is an example of reckless automation. Armed with a new data-driven project and $3.8 million dollars in federal funding, the police department experimented with a program that led to numerous "Custom Notification Plan" interventions: visits to the homes of persons identified as high risk.[61] These visits were formally described as opportunities to connect high risk persons to social service programs, but in many cases may have been no more than "going door-to-door notifying potential criminals not to commit any violent crimes."[62]

Consider another example: the mistaken arrest of Robert Julian-Borchak Williams.[63] Detroit police had been investigating the theft of $3800 worth of watches from a local store in 2018. An examiner for the Michigan state police uploaded a still image from the store's surveillance video to the state's facial recognition database. The system would have searched for potential matches in a database of 49 million photos.

The facial recognition technology used in this investigation was supplied by a private company that began as a mugshot management software company, which then added

---

[58] Id. at 8.

[59] Id.

[60] Id. at 1.

[61] Id. at 3; see also Andrew G. Ferguson, The Rise of Big Data Policing 38 (2017)(noting these visits involve "a home visit, usually by a senior police officer, a social worker, and a member of the community . . . During the visit, police hand deliver a 'custom notification letter' detailing what the police know about the individual's criminal past, as well as a warning about the future").

[62] Adrienne Balow & Judy Wang, CPD launches new "custom notifications" anti-violence program, WGN9, July 19, 2013, at https://wgntv.com/news/cpd-launches-new-custom-notifications-anti-violence-program/

[63] The facts here are taken extensively from Kashmir Hill, Wrongfully Accused by an Algorithm, N.Y. Times, Aug. 3, 2020, at https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.

facial recognition tools developed by subcontractors.[64] But the private company that contracts with the state does not measure these systems for accuracy or bias.

In Williams's case, the Detroit police would have received a report with potential matches generated by the software. Algorithms incorporated into the software used in the case were also found in a 2019 a federal study to have significant inaccuracies in identifying African-American and Asian faces as compared to Caucasian ones.[65] A store employee identified Williams from a photo lineup generated from that report.

Confronted with the video still, Williams asked "You think all black men look alike?"[66] Detroit detectives eventually acknowledged that they arrested the wrong person. But that admission came after Williams had been handcuffed and arrested at home in front of his wife and daughter, had his mug shot, fingerprints, and DNA sample taken, and had been held overnight in jail in 2020.[67] Williams subsequently sued the Detroit police for his wrongful arrest.[68]

The pivotal role of facial recognition in this widely publicized wrongful arrest will be discussed with the terms of algorithmic accountability, but it is also an example of reckless automation that harms the principles of respect for persons, beneficence, and justice. Should the police in Michigan have decided to use an automated system which was demonstrated to make racially disproportionate mistakes, and thus harms, to people? Does such a decision respect the autonomy of persons affected? Can we conclude that such a decision demonstrates attention to "fair procedures and outcomes" affecting the groups and individuals involved?[69] Was there special attention to the fact that any potential harms and mistakes would affect racial minorities disproportionately?[70]

---

[64] See id.

[65] See id., citing National Institute of Standards and Technology, Face Recognition Vendor Test (Dec. 2019), at https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf.

[66] See id.

[67] See id.

[68] Drew Harwell, Wrongfully arrested man sues Detroit police over false facial recognition match, Wash. Post, April 13, 2021, at https://www.washingtonpost.com/technology/2021/04/13/facial-recognition-false-arrest-lawsuit/

[69] Belmont Report, supra note xx, at 9.

[70] Belmont Report, supra note xx, at 9 ("One special instance of injustice results from the involvement of vulnerable subjects. Certain groups, such as racial minorities, the economically disadvantaged, the very

Conclusion

The growing calls for algorithmic accountability in the tools of artificial intelligence merit the attention they received. These calls for attention to bias, secrecy, and inaccuracy have special importance in the use of artificial intelligence in policing, where mistakes and biases can impose life-altering consequences. This essay, however, has offered a different perspective: characterizing some adoptions of artificial intelligence in policing as reckless automation, and one that might benefit from the conversations in human subjects research ethics. The long history of ethical lapses in human subjects research has prompted a robust framework that asks fundamental questions about individual consent, community impact, minimizing harms, and special attention to racial minorities, among other groups. The algorithmic accountability conversation brings an important perspective about *technological processes* to policing. This essay urges that a bioethical perspective can offer a perspective on the *human impacts* of policing automation.

sick, and the institutionalized may continually be sought as research subjects, owing to their ready availability in settings where research is conducted. Given their dependent status and their frequently compromised capacity for free consent, they should be protected against the danger of being involved in research solely for administrative convenience, or because they are easy to manipulate as a result of their illness or socioeconomic condition.").