

DIRTY DATA, BAD PREDICTIONS: HOW CIVIL RIGHTS VIOLATIONS IMPACT POLICE DATA, PREDICTIVE POLICING SYSTEMS, AND JUSTICE

RASHIDA RICHARDSON,* JASON SCHULTZ,** AND KATE CRAWFORD***

Law enforcement agencies are increasingly using algorithmic predictive policing systems to forecast criminal activity and allocate police resources. Yet in numerous jurisdictions, these systems are built on data produced within the context of flawed, racially fraught and sometimes unlawful practices ('dirty policing'). This can include systemic data manipulation, falsifying police reports, unlawful use of force, planted evidence, and unconstitutional searches. These policing practices shape the environment and the methodology by which data is created, which leads to inaccuracies, skews, and forms of systemic bias embedded in the data ('dirty data'). Predictive policing systems informed by such data cannot escape the legacy of unlawful or biased policing practices that they are built on. Nor do claims by predictive policing vendors that these systems provide greater objectivity, transparency, or accountability hold up. While some systems offer the ability to see the algorithms used and even occasionally access to the data itself, there is no evidence to suggest that vendors independently or adequately assess the impact that unlawful and bias policing practices have on their systems, or otherwise assess how broader societal biases may affect their systems.

In our research, we examine the implications of using dirty data with predictive policing, and look at jurisdictions that (1) have utilized predictive policing systems and (2) have done so while under government commission investigations or federal court monitored settlements, consent decrees, or memoranda of agreement stemming from corrupt, racially biased, or otherwise illegal policing practices. In particular, we examine the link between unlawful and biased police practices and the data used to train or implement these systems across thirteen case studies. We highlight three of these: (1) Chicago, an example of where dirty data was ingested directly into the city's predictive system; (2) New Orleans, an example where the extensive evidence of dirty policing practices suggests an extremely high risk that dirty data was or will be used in any predictive policing application, and (3) Maricopa County where despite extensive evidence of dirty policing practices, lack of transparency and public accountability surrounding predictive policing inhibits the public from assessing the risks of dirty data within such systems. The implications of these findings have widespread ramifications for predictive policing writ large. Deploying predictive policing systems in jurisdictions with extensive histories of unlawful police practices presents elevated risks that dirty data will lead to flawed, biased, and unlawful predictions which in turn risk perpetuating additional harm via feedback loops throughout the criminal justice system. Thus, for any jurisdiction where police have been found to engage in such practices, the use of predictive policing in any context must be treated with skepticism and mechanisms for the public to examine and reject such systems are imperative.

* Director of Policy Research, AI Now Institute at NYU.

** Professor of Clinical Law, NYU; Law & Policy Area Lead, AI Now Institute at NYU.

*** Distinguished Research Professor, New York University; Principal Researcher, Microsoft Research; Co-Founder, AI Now Institute at NYU. We wish to thank Andrew Ferguson, Barry Friedman, Meredith Whittaker, and Vincent Southerland for their invaluable insights and feedback on this essay. We also wish to thank Rebecca Sobel for her research assistance.

I. INTRODUCTION	3
A. <i>Dirty Data and “Juked Stats”</i>	
B. <i>What is Predictive Policing?</i>	
II. PREDICTIVE POLICING’S RELIANCE ON FLAWED DATA	10
A. <i>Overview of the Study</i>	
B. <i>Case Study 1: Philadelphia</i>	
C. <i>Case Study 2: New Orleans</i>	
D. <i>Case Study 3: Maricopa County</i>	
III. THE PROBLEMS OF USING OF POLICE DATA AS CURRENTLY CONSTITUTED	20
A. <i>Confirmation Feedback Loop</i>	
B. <i>The Boundless Reach of Bias Police Data</i>	
IV. CONCLUSION	25

I INTRODUCTION

A. Dirty Data and “Juked Stats”

In the crime TV series, *The Wire*, police are regularly instructed to shift arrest patterns and manipulate crime statistics in order to give the appearance of crime reduction. This is a practice called “juking the stats.” “If the felony rate won’t fall, you most certainly will,” one officer is instructed. The warning continues, “It’s Baltimore, gentlemen: the gods will not save you.”¹ Indeed, in Baltimore itself, the last ten years have resulted in federal investigations into systemic data manipulation, police corruption, falsifying police reports, and violence, including robbing residents, planting evidence, extortion, unconstitutional searches and other corrupt practices that result in innocent people being sent to jail.² As of 2018, Baltimore faces over fifty-five potential lawsuits in connection to police corruption and tainted data records.³

While *The Wire* fictionalized Baltimore’s very real history of police scandals, similar practices have been evidenced around the country. For example, in New York, “[m]ore than a hundred retired New York Police Department captains and higher-ranking officers said in a survey that the intense pressure to produce annual crime reductions led some supervisors and precinct commanders to manipulate crime statistics,” according to criminologists studying the department.⁴ These data manipulation practices were revealed in great detail through a survey and later a book published by two criminal justice professors, one of whom was a retired NYPD captain, on the controversial crime statistics reporting program known as CompStat.⁵ The survey, book, and subsequent media reports revealed incidents of precinct commanders going to a crime scene to persuade victims to not file complaints in order to artificially reduce serious crime statistics, as well as officers engaging in corrupt practices like planting drugs on innocent people or falsifying records in order to meet arrests and summons quotas.⁶ The logic and goals of these practices were to keep numbers for measured “serious crimes” low, since that data had to be reported to the FBI,

¹ *The Wire: Dead Soldiers* (HBO television broadcast Oct. 3, 2004).

² See CIVIL RIGHTS DIV., U.S. DEP’T OF JUSTICE, INVESTIGATION OF THE BALTIMORE CITY POLICE DEPARTMENT, 5–10 (2016), <https://www.justice.gov/crt/file/883296/download>; Luke Broadwater, *1,700 Cases Affected by Corrupt Baltimore Police Gun Trace Task Force, State’s Attorney Mosby Says*, BALT. SUN (June 5, 2018), <https://www.baltimoresun.com/news/maryland/crime/bs-md-ci-mosby-gttf-20180605-story.html>.

³ Justin Fenton, *‘Everything on the Table’ as Commission Begins Examining Corrupt Baltimore Police Gun Trace Task Force*, BALT. SUN (Oct. 16, 2018), <https://www.baltimoresun.com/news/maryland/crime/bs-md-ci-gttf-commission-first-meeting-20181016-story.html>; Joy Lepola, *Baltimore Faces Up to 55 Possible Lawsuits Over Police Corruption*, FOX 45 NEWS (June 19, 2018), <https://foxbaltimore.com/features/operation-crime-justice/baltimore-faces-up-to-55-possible-lawsuits-over-police-corruption>.

⁴ 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>.

⁵ JOHN A. ETERNO & ELI B. SILVERMAN, *THE CRIME NUMBERS GAME: MANAGEMENT BY MANIPULATION* (2012).

⁶ See, e.g., John Marzulli, *We Fabricated Drug Charges Against Innocent People to Meet Arrest Quotas, Former Detective Testifies*, N.Y. DAILY NEWS (Oct. 13, 2011), <https://www.nydailynews.com/news/crime/fabricated-drug-charges-innocent-people-meet-arrest-quotas-detective-testifies-article-1.963021>; Rashbaum, *supra* note 4; Graham Rayman, *The NYPD Tapes: Inside Bed-Stuy’s 81st Precinct*, VILLAGE VOICE (May 4, 2010), <https://www.villagevoice.com/2010/05/04/the-nypd-tapes-inside-bed-stuys-81st-precinct/>; Dan Rosenblum, *‘Quotas Exist’: Former NYPD Officers Discuss the Data-Driven Department*, POLITICO (May 4, 2012, 4:50 PM), <https://www.politico.com/states/new-york/city-hall/story/2012/05/quotas-exist-former-nypd-officers-discuss-the-data-driven-department-067223>.

while keeping the number of street arrests and summonses high, to give the appearance of community control.⁷

These “juked stats” were later confirmed by an NYPD-commissioned independent audit of the CompStat program,⁸ and the underlying practices were subsequently challenged in a series of lawsuits against the NYPD. The lawsuits resulted in a federal court opinion of almost two hundred pages finding that the NYPD engaged in over a decade of unconstitutional and racially biased practices and policies, which required systemic reforms and monitoring by a federal court for compliance.⁹ Similarly, Baltimore’s decades of police corruption resulted in a consent decree, signed in April 2017,¹⁰ between the City of Baltimore and the Department of Justice to address and reform police corruption and other unlawful practices. These court-based interventions primarily focus on cleaning up specific unconstitutional and corrupt processes and practices. But once the specified reforms are met there is little or no consideration of the need to address the police data that remains as an artifact of the prior unlawful conduct, and continues to shape predictive policing software going forward. So, in this Article, we ask a different question: What happens to predictive policing systems when police data records are influenced by systemic attempts to falsify crimes, plant evidence, and other actions that produce dirty data? In the absence of standardized data collection practices, how often do police departments or police technology vendors independently validate police records for accuracy or bias? How often is dirty data being included as the ground truth influencing predictive policing systems and other actors throughout the criminal justice system? What other forms of suspect or manipulated data is being ingested by predictive policing systems and how might this skew the predictions and subsequent recommendations? Can dirty data be remedied for subsequent use or is there a deeper and insurmountable problem derived from police practices and policies?

“Dirty data” is a term commonly used in the data mining research community to refer to “missing data, wrong data, and non-standard representations of the same data.”¹¹ For the purposes of this paper, we are expanding the term “dirty data” to include a new category that reflects the culture of data production in policing. This new category includes data that is derived from or influenced by corrupt, biased and unlawful practices, including data that has been intentionally manipulated or “juked,” as well as data that is distorted by individual and societal biases. Dirty data—as we use the term here—also includes data generated from the arrest of innocent people who had evidence planted on them or were otherwise falsely accused, in addition to calls for service or incident reports that reflect false claims of criminal activity.¹² In addition, dirty data incorporates subsequent uses that further distort police records, such as the systemic manipulation of crime statistics to try to promote particular public relations, funding, or political outcomes. Importantly, data can be subject to multiple forms of manipulation at once, which makes it

⁷ Rosenblum, *supra* note 6.

⁸ See DAVID N. KELLEY & SHARON L. MCCARTHY, THE REPORT OF THE CRIME REPORTING REVIEW COMMITTEE TO COMMISSIONER RAYMOND W. KELLY CONCERNING COMPSTAT AUDITING, 4–6 (2013), http://www.nyc.gov/html/nypd/downloads/pdf/public_information/crime_reporting_review_committee_final_report_2013.pdf

⁹ *Floyd v. City of New York*, 959 F. Supp. 2d 540 (2013).

¹⁰ *United States v. Baltimore Police Department*, 249 F. Supp. 3d 816 (2017); Consent Decree, *United States v. Police Department of Baltimore*, Case 1:17-cv-00099-JKB (Jan. 12, 2017), <https://www.justice.gov/opa/file/925056/download>.

¹¹ Won Kim et al., *A Taxonomy of Dirty Data*, 7 DATA MINING & KNOWLEDGE DISCOVERY 81 (2003).

¹² See *infra* Section III.A.

extremely difficult, if not impossible, for systems trained on this data to detect and separate “good” data from “bad” data, especially when the data production process itself is suspect. This challenge is notable considering that some prominent predictive policing vendors assume that the problems of “dirty data” in policing can be isolated and repaired through classic mathematical, technological, or statistical techniques.¹³

For example, in 2015 the NYPD entered a contract for a predictive policing system that would use the NYPD’s historical crime data to predict where crime is likely to occur in the future in order to help precincts determine where to dispatch officers.¹⁴ In June 2018, the Baltimore Police Department (BPD) expressed interest in acquiring a predictive policing system and other systems using police data.¹⁵ However, to date, neither the NYPD, BPD, or any of their technology vendors have clarified how they intend to address the “dirty data” problem that these systems may have. Given that these systems are shaped in large part by prior policing patterns, often reinforcing already known or ingrained biases, such risks are likely substantial.¹⁶ As an opinion piece in the *Baltimore Sun* noted, “Deploying officers based on crime statistics will simply return them to where they concentrate their time. As a result, the data often push officers into the same over-policed and over-criminalized communities.”¹⁷ This becomes part of what is known as the “bias in, bias out” phenomenon of predictive systems.¹⁸

This problem is further intensified when policing data is tainted by corruption and other unconstitutional or unethical police practices. To demonstrate this, we identified thirteen jurisdictions that are using or previously used predictive policing systems while also being subject to government investigations, consent decrees, or having been shown to be engaged in corrupt, racially biased, or otherwise illegal police practices. We then compared the evidence from the Department of Justice investigations and federal court adjudications with publicly available

¹³ P. Jeffrey Brantingham, *The Logic of Data Bias and Its Impact on Place-Based Predictive Policing*, 15 OHIO ST. J. CRIM L. 473, 485 (2018) (“The conclusion is that we need to work hard to figure out how to detect and correct for biases in police data rather than rejecting such data out of hand or accepting it without further thought.”).

¹⁴ Laura Nahmias & Miranda Neubauer, *NYPD Testing Crime-Forecast Software*, POLITICO (July 8, 2015, 5:52 AM), <https://www.politico.com/states/new-york/city-hall/story/2015/07/nypd-testing-crime-forecast-software-090820>; N.Y.C. POLICE DEP’T, DEVELOPING THE NYPD’S INFORMATION TECHNOLOGY, <http://www.nyc.gov/html/nypd/html/home/POA/pdf/Technology.pdf>.

¹⁵ NAT’L POLICE FOUND. & BALT. POLICE DEP’T, BALTIMORE POLICE DEPARTMENT TECHNOLOGY RESOURCE STUDY (2018), https://www.baltimorepolice.org/sites/default/files/General%20Website%20PDFs/BPD_Final_Technology_Inventor_y_Study_06-21-18.pdf; see also Justin Fenton, *5 Shot in Separate Incidents Friday in Baltimore, Police Say*, BALTIMORE SUN (Nov. 24, 2018, 10:50 AM), <https://www.baltimoresun.com/topic/crime-law-justice/law-enforcement/baltimore-police-department-ORGOV000345-topic.html>; Caroline Haskins, *Predictive Policing Tool’s Website Exposes Login Pages for 17 US Police Departments*, VICE (Oct. 30, 2018, 1:33 PM), https://motherboard.vice.com/en_us/article/wj9v9q/predictive-policing-tools-website-exposes-login-pages-for-17-us-police-departments (describing Baltimore as a city with a PredPol login, which suggests former or future use of the service).

¹⁶ See generally MICHELLE ALEXANDER, THE NEW JIM CROW: MASS INCARCERATION IN THE AGE OF COLORBLINDNESS (rev. ed. 2012); Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1146–51 (2017); Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 55–59 (2014); Ezekiel Edwards, *Predictive Policing Software Is More Accurate at Predicting Policing than Predicting Crime*, HUFFINGTON POST (Oct. 31, 2016), http://www.huffingtonpost.com/entry/predictive-policing-reform_us_57c6ffe0e4b0e60d31dc9120.

¹⁷ Michael Pinard, *Predicting More Biased Policing in Baltimore*, BALTIMORE SUN (Apr. 10, 2018, 10:20 AM), <https://www.baltimoresun.com/news/opinion/oped/bs-ed-op-0411-predictive-policing-20180410-story.html>.

¹⁸ See, e.g., Sandra G. Mayson, *Bias In, Bias Out*, 128 YALE L.J. (forthcoming 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3257004.

information regarding the jurisdiction's use of predictive policing systems to determine whether police data used to train or implement those systems was generated during the periods of unlawful and biased police activity.

Our analysis revealed nine jurisdictions where predictive policing systems appear to have used police data generated during periods when the department was found to have engaged in various forms of unlawful and biased police practices. This strongly indicates that these predictive systems are at high risk of being guided by tainted data, and the records of illegal and unconstitutional police practices. Below, we highlight three of the nine case studies to show both the direct and indirect linkages between dirty data and predictive policing as well as the dangers of obfuscating or inhibiting public conversation and accountability for how these systems are build and used.

B. What Is Predictive Policing?

Predictive policing generally describes any system that analyzes available data to predict either *where* a crime may occur in a given time window (place-based) or *who* will be involved in a crime as either victim or perpetrator (person-based). It is the latest iteration of data-driven crime analysis techniques that law enforcement agencies are increasingly relying on for crime control and forecasting. Few predictive policing vendors are fully transparent about how their systems operate, what specific data is used in each jurisdiction that deploys the technology, or what accountability measures the vendor employs in each jurisdiction to address potential accuracy, bias, or evidence of misconduct. Despite these looming questions, one known fact is that historical police data is the primary data source used to inform these systems, and while the specific data categories will vary by system, it can include information on past crimes (type of crime, time, and location), arrests and calls for service.¹⁹

Though it is too commonly assumed that police data is objective, it is embedded with political, social, and other biases.²⁰ Indeed, police data is a reflection of the department's practices and priorities, local, state or federal interests, as well as institutional and individual biases.²¹ In fact, even calling this information, "data," could be considered a misnomer, since "data" implies some type of consistent scientific measurement or approach.²² 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.²³ This lack of rigorous methodology and

¹⁹ See, e.g., Andrew Guthrie Ferguson, *The Police Are Using Computer Algorithms to Tell If You're a Threat*, TIME (Oct. 3, 2017), <http://time.com/4966125/police-departments-algorithms-chicago/>; AZAVEA, HUNCHLAB: UNDER THE HOOD (2015), <https://cdn.azavea.com/pdfs/hunchlab/HunchLab-Under-the-Hood.pdf>.

²⁰ See ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT (2017); Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 82 AM. SOC. REV. 977, 997–1000 (2017); Brian Jordan Jefferson, *Predictable Policing: Predictive Crime Mapping and Geographies of Policing and Race*, 108 ANNALS AM. ASS'N GEOGRAPHERS 1 (2018).

²¹ See FERGUSON, *supra* note 20; Barbara E. Armacost, *Organizational Culture and Police Misconduct*, 72 GEO. WASH. L. REV. 453, 494 (2004) (describing how the organizational culture in police departments affects misconduct and other unlawful practices); Kristian Lum & William Isaac, *To Predict and Serve?*, SIGNIFICANCE, Oct. 2016, at 14, <https://tss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x>.

²² See "RAW DATA" IS AN OXYMORON, 1-9 (Lisa Gitelman ed., 2013).

²³ See *id.*; Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827, 1855–65 (2015); Maria Ponomarenko & Barry Friedman, *Benefit-Cost Analysis of Public Safety: Facing the Methodological*

accountability leaves significant room for subjectivity in how or what officers choose to record about their activities, and few incentives for police leadership or other government actors to interrogate or validate records for accuracy, bias, misconduct, or to identify proactive reforms.²⁴

Moreover, there is no evidence of predictive policing vendors independently validating the police records used to develop their systems or within the jurisdictions where the technology is deployed. Instead, even those vendors willing to acknowledge biased data as a problem merely attempt to isolate or segregate it from what is presumably “clean” data instead of seeing it as an indicator of the potential unreliability of the entire data set from that jurisdiction. For example, the place-based predictive policing company, PredPol, claims that it primarily collects and analyzes so-called “victim data” – time of date, location of crime and type of crime from reports.²⁵ This is an attempt to position such data as objective and untainted. On the other hand, PredPol excludes drug-related offenses (which have known and well-documented racial disparities) and traffic citation data from its predictions “to remove officer bias from the equation and eliminate the risk of generating predictions based on officer discretion.”²⁶ Yet drug-related offenses and traffic stops are hardly the only crimes where officers exercise bias or discretion, let alone the only crimes where corruption, discrimination, racial profiling, or other dirty police practices exist. For example, even deciding which circumstances to investigate, define, and document as a “crime” can be a matter of officer discretion, as our case studies show.

Moreover, examining the context and use of policing practices to generate data is also important because encounters with police are the most common point of entry for individuals into the criminal justice system. Many subsequent decision-making processes in criminal justice, including those during prosecution, pretrial services, adjudication, sentencing, parole and corrections, derive their analysis from policing data inputs.²⁷ As such, an examination of policing practices and the data that acts as a record of such practices is required for any informed discussion of the possible harms and benefits of predictive policing systems whose conclusions are shaped by such data. Yet, despite this need for scrutiny, policing is often the most funded and least regulated of all the government agencies.²⁸ A driver of this imbalance is that law enforcement agencies are given undue deference by all branches of government.²⁹

Challenges, 8 J. BENEFIT-COST ANALYSIS 305 (2017); Marvin E. Wolfgang, *Uniform Crime Reports: A Critical Appraisal*, 111 U. PA. L. REV. 708 (1963) (detailing the flaws in the FBI’s Uniform Crime Reporting).

²⁴ Rachel Harmon, *Why Do we (Still) Lack Data on Policing?*, 96 MARQ. L. REV. 1119, 1129–33 (2013) (describing how cities and police departments actively inhibit the collection of information about police, and how federal efforts of data collection and review are stymied by institutional and legal limitations).

²⁵ *Machine Learning and Policing*, PREDPOL: PREDICTIVE POLICING BLOG (July 19, 2017, 10:00 AM), <http://blog.predpol.com/machine-learning-and-policing>.

²⁶ *Id.*

²⁷ See JENNIFER A. TALLON ET AL., CTR. FOR COURT INNOVATION, THE INTELLIGENCE-DRIVEN PROSECUTION MODEL: A CASE STUDY IN THE NEW YORK COUNTY DISTRICT ATTORNEY’S OFFICE (2016), https://www.courtinnovation.org/sites/default/files/documents/IDPM_Research_Report_FINAL.PDF (describing prosecutorial use of police data); NORTHPOINTE, PRACTITIONERS GUIDE TO COMPAS (2012), http://www.northpointeinc.com/files/technical_documents/FieldGuide2_081412.pdf (describing the use of police data for pre-trial, sentencing, parole, and probation decisions); MHS ASSESSMENTS, LEVEL OF SERVICE/CASE MANAGEMENT INVENTORY (2004), https://www.assessments.com/assessments_documentation/LSCMI_Tech_Brochure.pdf (describing the use of police data in parole case management).

²⁸ Friedman & Ponomarenko, *supra* note 23 (describing how police remain the least regulated among government bodies by conventional checks and balances of government).

²⁹ See *id.*; Ponomarenko & Friedman, *supra* note 23.

It is a common fallacy that police data is objective and reflects actual criminal behavior, patterns, or other indicators of concern to public safety in a given jurisdiction. In reality, police data reflects the practices, policies, biases, and political and financial accounting needs of a given department.³⁰ For instance, some data relevant to crime patterns and public safety indicators, such as drug use or police misconduct, are not reflected in police data or available publicly, which reflects underlying forms of political accounting and public relations.³¹ Hence, actual crime data is often incomplete or distorted. The Department of Justice has estimated that less than half of violent crimes and even fewer household property crimes are reported to the police.³² The type of criminal activity recorded by police also depends on which law enforcement agency has jurisdiction over which crimes.³³ Research also suggests that when a group feels less favorable toward local police, they are less likely to report a crime they witness.³⁴ Even when reported, errors and bias in how the police record reported crimes result in distorted data. The Los Angeles Police Department, for example, misrecorded a staggering 14,000 serious assaults as minor offenses from 2005 to 2012. This error was not discovered until 2015, by which time the LAPD had already begun its work with the predictive policing company, PredPol.³⁵

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.³⁶ 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 drug-related arrest data, from data categories that are customarily considered objective, such as calls for service data.³⁷ Moreover, even where such distinctions are possible, they would have to occur on

³⁰ See Armacost, *supra* note 21; Rachel Harmon, *Promoting Civil Rights Through Proactive Policing Reform*, 62 STAN. L. REV. 1 (2009).

³¹ See Armacost, *supra* note 21; Harmon, *supra* note 30.

³² LYNN LANGTON ET AL., U.S. DEP'T OF JUSTICE, VICTIMIZATION NOT REPORTED TO THE POLICE, 2006-2010 (2012).

³³ For example, the White Collar Crime Risk Zones Map satirically shows what predictive policing might look like if it were based on the data in the Financial Industry Regulatory Authority (FINRA) database of financial crimes instead of the NYPD database. Brian Clifton, Sam Lavigne & Francis Tseng, *White Collar Crime Risk Zones*, 59 NEW INQUIRY (Mar. 2017), <https://whitecollar.thenewinquiry.com>.

³⁴ See EMILY EKINS, CATO INST., POLICING IN AMERICA: UNDERSTANDING PUBLIC ATTITUDES TOWARD THE POLICE. RESULTS FROM A NATIONAL SURVEY (2016); Lum & Isaac, *supra* note 21 (“Bias in police records can also be attributed to levels of community trust in police, and the desired amount of local policing These effects manifest as unequal crime reporting rates throughout a precinct.”).

³⁵ Ben Poston, Joel Rubin, & Anthony Pesce, *LAPD Underreported Serious Assaults, Skewing Crime Stats for 8 Years*, L.A. TIMES (Oct. 15, 2015, 4:47 AM), <https://www.latimes.com/local/cityhall/la-me-crime-stats-20151015-story.html>.

³⁶ For example, a guide on the place-based predictive policing system, HunchLab, notes that bias in the enforcement of certain crimes can distort police data used in its system. As a corrective measure, it cites that “[f]or quality of life type crimes, we tend to use records that reflect the public’s call for services, which does not suffer from an enforcement bias.” Yet, the company fails to acknowledge or account for the fact that the public’s call for services can reflect other societal bias that can also distort the data, as discussed below in Section III.A. HUNCHLAB, A CITIZEN’S GUIDE TO HUNCHLAB 26 (2017), <http://robertbrauneis.net/algorithms/HunchLabACitizensGuide.pdf>.

³⁷ Harmon, *supra* note 24, at 1129–33 (“Though officers will collect information when police chiefs and local governments require them to do so, they will collect only that information and only in the form mandated. . . . Other times, police departments simply fail to produce records that could improve political and regulatory decision-making about intrusive police activities.”); see also COMM. ON PROACTIVE POLICING: EFFECTS ON CRIME, CMTYS., & CIVIL LIBERTIES, NAT’L ACADS. OF SCIS., ENG’G, & MED., PROACTIVE POLICING: EFFECTS ON CRIME AND COMMUNITIES

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.³⁸ There is a dearth of objective and comprehensive analysis of the efficacy and impact of certain police practices or policies, coupled with an unwillingness or inability on the part of police departments to investigate and monitor themselves impartially.³⁹ To overcome the challenge that this information deficit poses, our study relied on the findings of government-commissioned investigations, federal court monitored settlements, consent decrees, or memoranda of agreement. These types of agreements “begin with investigations of allegations of systemic police misconduct and, when the allegations are substantiated, end with comprehensive agreements designed to support constitutional and effective policing and restore trust between police and communities.”⁴⁰ Additionally, the findings related to these investigations, agreements, and federal court adjudications are generally considered substantiated and irrefutable, despite the fact that the settlements involve no finding of guilt. The required reforms typically prohibit the identified problematic practices, and a federal judge or third-party monitor has authority to enforce compliance with such prohibitions.⁴¹ Given this, we conclude that the data generated during the time periods covered by these findings sufficiently reflects bias and misassumptions, at the very least, embedded within police practices.

There is significant research and litigation raising concerns of bias in policing and the broader criminal justice system, but much of this scrutiny is focused on specific actors, practices or newly adopted systems.⁴² Complimenting these approaches, we examine how individual and collective practices by actors in and outside of the criminal justice system are reflected in the data that is generated and subsequently used throughout the criminal justice system, often without adequate transparency, accountability, oversight, or public engagement. We argue that biased and dirty data in the criminal justice system is systemic pervasive; therefore, solutions, especially technological ones, can neither erase dirty data, nor presume other policing data from the same jurisdiction is unproblematic.

292 (David Weisburd & Malay K. Majmudar, eds., 2018), <https://www.nap.edu/read/24928/chapter/9#292> (describing a study that leaves open the question of whether calls for service are racially biased).

³⁸ Harmon, *supra* note 24, at 1129 (“Even when departments collect information, they may do so in ways that make it impossible to aggregate the records or compare them with data from other departments.”).

³⁹ See Ponomarenko & Friedman, *supra* note 23.

⁴⁰ CIVIL RIGHTS DIV., U.S. DEP’T OF JUSTICE, THE CIVIL RIGHTS DIVISION’S PATTERN AND PRACTICE POLICE REFORM WORK: 1994-PRESENT 1 (2017).

⁴¹ Peter M. Shane, *Federal Policy Making by Consent Decree: An Analysis of Agency and Judicial Discretion*, 1987 U. CHI. LEGAL F. 241.

⁴² See, e.g., Mark W. Bennett, *Implicit Racial Bias in Sentencing: The Next Frontier*, 126 YALE L.J. F. 391 (2017); Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109 (2017); NOAH KATZ ET AL., UCLA LABOR CTR., GET TO WORK OR GO TO JAIL (2016), <https://www.labor.ucla.edu/publication/get-to-work-or-go-to-jail/> (describing how probation and parole conditions force low-income people to choose between bad or potentially unpaid jobs and jail time); ACLU OF OR., ROADBLOCKS TO REFORM: DISTRICT ATTORNEYS, ELECTIONS, AND THE CRIMINAL JUSTICE STATUS QUO (2016), http://aclu-or.org/sites/default/files/Roadblocks_to_Reform_Report_ACLUOR.pdf (detailing how district attorneys play a significant role in blocking progressive criminal justice reform and maintain the status quo out of self-interest).

II PREDICTIVE POLICING'S RELIANCE ON FLAWED DATA

A. Overview of Study

For this Article, we identified thirteen jurisdictions that are currently using predictive policing systems or have previously engaged in predictive policing pilots or studies while under government commission investigations or obliged by federal court monitored settlements, consent decrees, or memoranda of agreement that found that the police departments engaged in corrupt, racially biased, or otherwise illegal police practices. These jurisdictions include Baltimore (MD), Boston (MA), Chicago (IL), Ferguson (MO), Miami (FL), Maricopa County (AZ), Milwaukee (WI), New Orleans (LA), New York (NY), Newark (NJ), Philadelphia (PA), Seattle (WA), and Suffolk County (NY).⁴³ We compared the substantiated evidence and other findings of unlawful or biased police practices from the Department of Justice investigations or federal court adjudications with publicly available information regarding the jurisdiction's use of predictive policing systems to determine whether the police data used to train or implement the predictive policing system(s) was generated during the periods of unlawful and biased police activity.

The lack of public transparency about policing and predictive policing systems often makes it difficult to draw a straight line between the dirty data produced by the police departments in the reviewed jurisdictions and the predictive policing systems deployed. However, we found evidence to suggest that nine jurisdictions where predictive policing systems used some form of dirty data generated during periods when the department was found to have engaged in forms of corrupt police practices. Evidence in four other jurisdictions was less dispositive.

The following three case studies highlight our thesis. Chicago is an example of a jurisdiction where we found strong evidence that suggests the predictive policing system was using dirty data. Second, New Orleans is an example of a jurisdiction where the extensive dirty policing practices suggest an extremely high likelihood that *any* predictive policing application would be heavily influenced by dirty data yet because the public has been blocked from proper transparency and accountability mechanisms, the extent of the problem is not fully known. Finally, Maricopa County is an example of a jurisdiction where extensive dirty policing practices suggest an extremely high risk of dirty data being used for predictive policing not only for the County itself, but also for adjacent jurisdictions where data or police resources are shared. Again, lack of public transparency and accountability inhibit a forensic examination of the risks. Maricopa County also demonstrates an emerging trend of local police departments engaging in immigration enforcement, which complicates many of the aforementioned issues regarding the constitutionality, transparency, accountability and oversight of police practices.

B. Case Study I: Chicago

The Chicago Police Department (CPD) has a lengthy and well-documented history of corrupt, abusive, and biased practices, dating back to a 1972 blue-ribbon panel finding of extreme police misconduct that disproportionately affected residents of color.⁴⁴ In the years since, there have been several notable investigations and legal challenges, including evidence of over one

⁴³ See *infra* Appendix.

⁴⁴ The Misuse of Police Authority in Chicago: A Report and Recommendations Based on Hearings before the Blue Ribbon Panel Convened by the Honorable Ralph H. Metcalfe, First Congressional District of Illinois (June 26, July 17, 24, 31, 1972), <https://chicagopatf.org/wp-content/uploads/2016/01/metcalfe-report-1972.pdf>.

hundred cases of CPD torturing Black men between 1972 and 1991⁴⁵ and a lawsuit challenging CPD's inequitable deployment of police to emergency calls in neighborhoods with higher minority populations.⁴⁶ Given this breadth of issues, the remainder of our discussion of CPD practices, policies and data is limited to the last decade.

In March 2015, the ACLU of Illinois issued a groundbreaking report detailing the CPD's fraught history of stop and frisk practices.⁴⁷ Using CPD records on stops that occurred in 2012 and 2013 and four months of contact card data⁴⁸ from 2014, the report concluded that a significant number of CPD stops and frisks were unlawful, and Black residents were disproportionately subjected to these unlawful stops.⁴⁹ The report, which also found significant deficiencies in CPD data and data collection practices,⁵⁰ led to a settlement agreement in August 2015 requiring ongoing independent evaluation of CPD practices and procedures, data collection, officer training, and reform of investigatory street stop practices.⁵¹ A former federal judge oversaw the agreement and published regular public reports assessing CPD's compliance with the agreement's requirements, which revealed CPD's continued engagement in unlawful practices and the data reflected significant race and gender bias.⁵²

During this same period, Chicago received national attention due to public outcry and city-wide protests following the release of a videotape showing the fatal shooting of Laquan McDonald, a Black 17-year-old, by a CPD officer. This led to the Illinois Attorney General's December 2015⁵³ request to the Department of Justice to investigate the CPD, which resulted in a yearlong investigation of CPD and the Independent Police Review Authority, the body responsible for

⁴⁵ [The Chicago Torture Archive, http://web.archive.org/web/20161028161512/http://chicagotorturearchive.uchicago.edu/.](http://web.archive.org/web/20161028161512/http://chicagotorturearchive.uchicago.edu/)

⁴⁶ Complaint, Cent. Austin Neighborhood Ass'n v. City of Chicago, 11 CH 37299 (Ill. Cir. Ct. Oct. 27, 2011), https://www.aclu-il.org/sites/default/files/field_documents/complaint.pdf (challenging the City of Chicago for failing to ensure police are deployed equitably across the City, which resulted in delayed police responses and high rates of serious violent crimes in neighborhoods with significant ethnic minority populations).

⁴⁷ ACLU OF ILL., STOP AND FRISK IN CHICAGO (2015), https://www.aclu-il.org/sites/default/files/wp-content/uploads/2015/03/ACLU_StopandFrisk_6.pdf.

⁴⁸ Contact cards are forms CPD officers filled out after street stops not resulting in tickets or arrests. *Id.* at 6.

⁴⁹ *Id.* at 6–10.

⁵⁰ *Id.* at 14–15 (“[T]he Chicago Police Department does not record stop and frisks in a way that reflects the full picture of what is happening on the streets of Chicago. Chicago does not have a single database of all stops available to the public and officers make no records of frisks.”).

⁵¹ Investigatory Stop and Protective Pat Down Settlement Agreement, City of Chicago, Chicago Police Department & ACLU of Illinois (Aug. 2015), <https://www.aclu-il.org/sites/default/files/wp-content/uploads/2015/08/2015-08-06-Investigatory-Stop-and-Protective-Pat-Down-Settlement-Agreeme....pdf>.

⁵² See ARLANDER KEYS, THE CONSULTANT'S FIRST SEMI-ANNUAL REPORT ON THE INVESTIGATORY STOP AND PROTECTIVE PAT DOWN AGREEMENT FOR THE PERIOD JANUARY 1, 2016 – JUNE 30, 2016 (Mar. 23, 2017), <https://www.aclu-il.org/sites/default/files/wysiwyg/the-consultants-first-semiannual-report-3-23-17.pdf>; ARLANDER KEYS, THE CONSULTANT'S SECOND SEMI-ANNUAL REPORT: INVESTIGATORY STOP & PROTECTIVE PAT DOWN AGREEMENT (MAR. 5, 2018), https://www.aclu-il.org/sites/default/files/field_documents/the_consultants_second_semi-annual_report.pdf.

⁵³ In the same month, Chicago Mayor Rahm Emanuel separately created a Police Accountability Task Force. *What is the Task Force*, CHI. POLICE ACCOUNTABILITY TASK FORCE, <https://chicagopatf.org/about/what-is-police-accountability-task-force/> (last visited Feb. 1, 2019). The Task Force subsequently released a report with more than 100 recommendations for reform in April 2016. See POLICE ACCOUNTABILITY TASK FORCE, RECOMMENDATIONS FOR REFORM: RESTORING TRUST BETWEEN THE CHICAGO POLICE AND THE COMMUNITIES THEY SERVE, RECOMMENDATIONS (Apr. 2016), <https://chicagopatf.org/wp-content/uploads/2016/04/PATF-Complete-Recommendations-.pdf>.

investigating police misconduct. For this investigation, federal officials reviewed CPD records between 2011 and 2016, performed local visits, and met with community members, City officials, CPD staff, and local unions.⁵⁴ A report on this investigation concluded that CPD engaged in a pattern or practice of unconstitutional use of force; poor data collection to identify and address unlawful conduct;⁵⁵ systemic deficiencies in training and supervision; systemic deficiencies in accountability systems that contribute to the pattern or practice of unconstitutional conduct; and unconstitutional conduct that disproportionately affects Black and Latino residents.⁵⁶

Despite these findings, the Department of Justice, under direction of former Attorney General Jeff Sessions, announced that it would not seek a consent decree to reform CPD.⁵⁷ This prompted several lawsuits against the City of Chicago to seek a consent decree to reform CPD that would address the findings and recommendations of the Department of Justice investigation.⁵⁸ The City of Chicago opted to negotiate a consent decree with only one plaintiff of the lawsuits, the Illinois Attorney General, and it entered a Memorandum of Agreement with the other plaintiffs, over a dozen community and civil rights organizations, agreeing to pause the lawsuits during the consent decree negotiations and providing the local organizations the right to object if the decree is inadequate.⁵⁹ In January 2019, a federal judge approved a consent decree between the Illinois Attorney General's Office and the City, which requires extensive reforms of CPD practices, policies, and oversight.⁶⁰

Amid these years of overlapping investigations and challenges of CPD practices and policies, CPD developed the Strategic Subject List (SSL), a computerized assessment tool that incorporates numerous sources of information to analyze crime as well as identifies and ranks

⁵⁴ U.S. DEP'T OF JUSTICE, CIVIL RIGHTS DIV. & U.S. ATTORNEY'S OFFICE, N. DIST. OF ILL., INVESTIGATION OF THE CHICAGO POLICE DEPARTMENT 2–3 (2017), <https://www.justice.gov/opa/file/925846/download>.

⁵⁵ The report described video or other evidence demonstrating unconstitutional conduct by CPD that was falsely or inaccurately described in reports provided by officers. In some instances, these inaccurate reports resulted in false arrests and convictions. The report concluded that the extent of CPD's unconstitutional conduct may be larger than CPD records indicate given the large volume of reported incidents that do not have exculpatory or corroborating evidence, which further demonstrates our concerns regarding "dirty data" generated by unlawful and bias police practices. *Id.* at 36–37, 74–79.

⁵⁶ *Id.* at 15, 144–145.

⁵⁷ United States' Statement of Interest Opposing Proposed Consent Decree, *Illinois v. City of Chicago*, No. 17-cv-6260 (N.D. Ill. Oct. 12, 2018), <https://www.documentcloud.org/documents/5001703-Trump-Chicago-police-reform.html>; see also Dan Hinkel, *U.S. Attorney General Says Chicago Police Consent Decree Should Be Tossed, While Activists Seek Tighter Rules*, CHI. TRIB. (Oct. 12, 2018), <https://www.chicagotribune.com/news/local/breaking/ct-met-consent-decree-chicago-police-shootings-20181012-story.html>.

⁵⁸ See Complaint, *Illinois v. City of Chicago*, No. 17-cv-6260 (N.D. Ill. Aug. 17, 2017), http://www.illinoisattorneygeneral.gov/pressroom/2017_08/City_of_Chicago_lawsuit_filedcomplaint.pdf; Amended Complaint, *Cmtys. United v. City of Chicago*, No. 17-cv-7151 (N.D. Ill. Nov. 28, 2017), https://www.aclu-il.org/sites/default/files/field_documents/amended_complaint_0.pdf.

⁵⁹ Memorandum of Agreement Between the Office of the Illinois Attorney General and the City of Chicago and *Campbell v. City of Chicago* Plaintiffs and *Communities United v. City of Chicago* Plaintiffs (Mar. 2018), https://www.aclu-il.org/sites/default/files/field_documents/executed_moa.pdf.

⁶⁰ Consent Decree, *Illinois v. City of Chicago*, No. 17-cv-6260 (N.D. Ill. Sept. 13, 2018), <http://chicagopoliceconsentdecree.org/wp-content/uploads/2018/09/Illinois-v.-Chicago-Final-Consent-Decree-with-signatures.pdf>; Dan Hinkel, *What will change about the Chicago police under the consent decree?*, CHI. TRIB. (Feb. 1, 2019), <https://www.chicagotribune.com/news/local/breaking/ct-met-cb-what-will-chicago-police-consent-decree-do-20190131-story.html>.

individuals at risk of becoming a victim or possible offender in a shooting or homicide.⁶¹ The tool was developed by the Illinois Institute of Technology and funded through the Department of Justice Bureau of Justice Assistance grant program, and some version of the tool has been used since 2012. When information on the SSL was first made publicly following FOIA litigation in 2017, the dataset included 398,684 individuals.⁶²

The SSL ranks and assigns risk tiers ranging from very low to very high⁶³ to individuals based on the following variables: the number of times an individual was a victim of a shooting; the individual's age during latest arrest; the number of times the individual was a victim of aggravated battery or assault; trend in criminal activity⁶⁴; number of prior arrest for unlawful use of a weapon; number of prior arrest for violent offenses; number of prior narcotics arrests; and gang affiliation⁶⁵.⁶⁶ It is notable that a majority of these variables are based on arrest records, rather than convictions, which not only means that people who have not committed crimes may end up on the list but the list likely reflects CPD's unlawful and bias practices.⁶⁷ These facts were both confirmed by analysis of the SSL dataset. Independent analysis by Upturn and *The New York Times* found that more than one third of individuals on this list have never been arrested or a victim of a crime, and almost seventy percent of that cohort received a high risk score.⁶⁸ The SSL data also

⁶¹ See *Strategic Subject List*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List/4aki-r3np> (last visited Feb. 3, 2019); Chicago Police Department: Subject Assessment and Information Dashboard (SAID), Special Order S09-11 (Jan. 9, 2019), <http://directives.chicagopolice.org/directives/data/a7a57b85-155e9f4b-50c15-5e9f-7742e3ac8b0ab2d3.html>; Chicago Police Department: Custom Notifications in Chicago, Special Order S10-05 (Oct. 6, 2015), <http://directives.chicagopolice.org/directives/data/a7a57bf0-1456faf9-bfa14-570a-a2deebf33c56ae59.html> [hereinafter Custom Notifications in Chicago].

⁶² *Strategic Subject List - Dashboard*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Dashboard/wgnt-sjgb> (last visited Feb. 3, 2019); see also Brianna Posadas, *How Strategic Is Chicago's "Strategic Subject List"?* Upturn Investigates., MEDIUM (June 22, 2017), <https://medium.com/equal-future/how-strategic-is-chicagos-strategic-subjects-list-upturn-investigates-9e5b4b235a7c>.

⁶³ When individuals are assessed as high risk, they are subject to heightened police scrutiny. Custom Notifications in Chicago, *supra* note 63.

⁶⁴ This is not defined and there is no information on how CPD calculates this number.

⁶⁵ This variable is included in the SSL dataset but while the Mayor's office claims that this variable is no longer used, the CPD confirmed its inclusion as a variable. See Yana Kunichoff & Patrick Sier, *The Contradictions of Chicago Police's Secretive List*, CHI. MAG. (Aug. 21, 2017), <https://www.chicagomag.com/city-life/August-2017/Chicago-Police-Strategic-Subject-List/>.

⁶⁶ See *Strategic Subject List - Dashboard*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Dashboard/wgnt-sjgb> (last visited Feb. 3, 2019).

⁶⁷ SSL data shows that the arrests of people identified on the list overlap with areas that are heavily targeted by CPD for patrol, which is documented through the contact cards police fill out after an investigatory street stop. The areas that are subject to heightened CPD presence and SSL enforcement are concentrated in the South and West sides of Chicago, which are predominately non-white and heavily low-income neighborhoods. See *Strategic Subject List - Dashboard*, CHI. DATA PORTAL, <https://data.cityofchicago.org/Public-Safety/Strategic-Subject-List-Dashboard/wgnt-sjgb> (last visited Feb. 3, 2019); *Map: Investigatory Stop Reports by District*, OFF. INSPECTOR GEN.: INFO. PORTAL (Aug. 20, 2018), <https://informationportal.igchicago.org/map-investigatory-stop-reports-by-beat-and-district/>; see also Heather Cherone, *Here's How Many Officers Are Patrolling Your Neighborhood*, DNAINFO: CHI. (Apr. 17, 2017), <https://www.dnainfo.com/chicago/20170417/logan-square/heres-how-many-officers-are-patrolling-your-neighborhood-watchdog/> ("Most officers are assigned to patrol the South and West Sides . . . [T]he Harrison and Englewood police districts on Chicago's West and South sides are home to about 5 percent of the city's population . . . Those two districts have the highest number of officers — 888 — assigned to patrol them, accounting for 12 percent of the total number of officers assigned to Chicago's police districts.").

⁶⁸ See 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->

revealed that fifty-six percent of black men under the age of 30 in Chicago have a risk score on the SSL, and this is the same population that has been disproportionately affected by CPD's unlawful and bias practices identified in the Department of Justice and ACLU reports.⁶⁹ These revelations are even more troubling in light of the conclusions of the only known validation study performed on an early version of the SSL by the RAND Corporation. The RAND study found the SSL was not successful in reducing gun violence or reducing the likelihood of victimization, inclusion on the SSL only had a direct effect on arrests, and the researchers noted these outcomes raised significant privacy and civil rights considerations.⁷⁰

There are also concerns regarding how the SSL predictions and risk scores are used by CPD in the field. A CPD internal directive and public statements claim that the SSL is used to target individuals with social services as part of the custom notification procedure, which is part of a citywide violence interventional model.⁷¹ Yet, the same CPD directive also encourages “[t]he highest possible charges” to be sought for any individuals on the SSL that received a custom notification and are subsequently arrested,⁷² and the RAND study observed that most police districts did not focus SSL enforcement on social service interventions.⁷³ CPD does not publicly release data on successful interventions, but available data and press coverage on CPD's SSL enforcement indicates arrests as a primary outcome, and in some cases a stated goal.⁷⁴ This

high-risk-list.html; Brianna Posadas, *How Strategic Is Chicago's "Strategic Subject List"? Upturn Investigates.*, MEDIUM (June 22, 2017), <https://medium.com/equal-future/how-strategic-is-chicagos-strategic-subjects-list-upturn-investigates-9e5b4b235a7c>. These findings repudiate prior CPD official statements suggesting that the SSL is populated with individuals that were known for driving gun violence, and reporting by the *Chicago Sun-Times* indicating that the list includes individuals that have previously been arrested and fingerprinted in Chicago since 2013. See Mick Dumke & Frank Main, *A Look Inside the Watch List Chicago Police Fought to Keep Secret*, CHI. SUN-TIMES (May 18, 2017), <https://chicago.suntimes.com/politics/what-gets-people-on-watch-list-chicago-police-fought-to-keep-secret-watchdogs/>.

⁶⁹ See Kunichoff & Sier, *supra* note 67; ACLU OF ILL., *supra* note 47, at 9; U.S. DEP'T OF JUSTICE, *supra* note 54, at 68, 143–50.

⁷⁰ Jessica Saunders et al., *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. EXPERIMENTAL CRIMINOLOGY 347, 366–67 (2016).

⁷¹ Custom Notifications in Chicago, *supra* note 63.; Mick Dumke & Frank Main, *A Look Inside the Watch List Chicago Police Fought to Keep Secret*, CHI. SUN-TIMES (May 18, 2017), <https://chicago.suntimes.com/politics/what-gets-people-on-watch-list-chicago-police-fought-to-keep-secret-watchdogs/> (“Lewin and other officials also say they don't rely on the scores alone when deciding which people to keep track of. Hundreds of people are flagged for interventions based on outstanding arrest warrants and ‘human intelligence’ in addition to their scores — so those with the highest scores aren't always the primary targets, Lewin says.”).

⁷² Custom Notifications in Chicago, *supra* note 63.

⁷³ Saunders et al., *supra* note 72, at 356 (noting district level guidance on SSL enforcement only occurred in 18.7% of COMPSTAT meetings observed and such guidance included “change the focus from arresting SSL subjects for minor offenses (for which they would be immediately released) to finding ways to detain SSL subjects over the long term”).

⁷⁴ See, e.g., Kunichoff & Sier, *supra* note 67 (“[I]n 2016, 1,024 notifications were attempted, 558 were completed, and only 26 people attended a call-in, where police officers, social workers, and others offer support services. . . . To put this in perspective, CPD has stated that 280 individuals with SSL scores were arrested in four gang raids during a six-month span in 2016.”); Jeremy Gerner, *In Crackdown on Violence, Chicago Police Arrest More Than 100 in Gang Raids*, CHI. TRIB. (May 20, 2016), <https://www.chicagotribune.com/news/local/breaking/ct-chicago-police-crackdown-on-violence-met-20160520-story.html> (“Chicago police have carried out an extensive gang takedown, arresting more than 115 people on the department's ‘strategic subject list’ — those believed to be most prone to violence.”); Sam Charles, *30 Arrested in Raids Aimed at Curbing Memorial Day Weekend Violence*, CHI. SUN-TIMES (May 27, 2017), <https://chicago.suntimes.com/news/30-arrested-in-raids-aimed-at-curbing-memorial-day-weekend-violence/amp/> (“All those taken into custody so far, along with those still being sought, are on the department's Strategic Subject List ‘Our goal was to identify the people that we think are driving the violence, that are going

disconnect between the SSL stated goals and outcomes was also discussed in the RAND study, which suggested the lack of a centralized crime prevention strategy and district-level guidance on how to use the SSL in the field may undermine any potential utility of the technology as a crime prevention strategy.⁷⁵ The study found that most CPD officers did not receive guidance on how to use the SSL predictions which resulted in officers merely increasing contacts with individuals on the list.⁷⁶ The RAND researchers noted “. . . it is not at all evident that contacting people at greater risk of being involved in violence—especially without further guidance on what to say to them or otherwise how to follow up—is this relevant strategy to reduce violence.”⁷⁷ This observation was also supported by the study’s results, which found increased police contacts with individuals on the SSL had no direct effect on arrest or victimization.⁷⁸

Despite these demonstrated concerns about the SSL and the underlying CPD practices and policies that still await reform, in 2017 the CPD entered a contract with the University of Chicago Crime Lab to develop and implement additional data-driven crimefighting strategies that will use predictive analytics and the SSL.⁷⁹ There is no evidence that this new initiative intends to account for ongoing consent decree negotiations or otherwise address the CPD’s unlawful and bias practices, including the dirty data generated by decades of these practices.

C. Case Study 2: New Orleans

The New Orleans Police Department (NOPD) has been investigated by the Department of Justice twice. The first investigation began in 1996 focusing on a wide-range of police misconduct, but it ended without a consent decree in 2004 because the NOPD pledged to reform itself.⁸⁰ In 2010 at the invitation of then-Mayor Mitchell J. Landrieu,⁸¹ the Department reopened its investigation of NOPD reviewing records between 2005 and 2011. The Department of Justice subsequently issued a report finding the NOPD engaged in a pattern or practice of excessive force disproportionately affecting Black residents; unlawful stops, searches and arrests; failure to detect, prevent, or address bias-based profiling and other discriminatory policing on basis of race, national origin, and LGBT status; racial disparities in arrest rates and other police data; and gender discrimination in the failure to adequately respond to and investigate violence against women.⁸²

to create the problems and snatch them and let them spend the weekend in Cook County Jail,’ [Organized Crime Bureau Chief Anthony Riccio] said.”)

⁷⁵ Saunders et al., *supra* note 72, at 355–56, 366–677.

⁷⁶ *Id.* at 363 (“[I]ncreased surveillance does appear to be caused by being placed on the SSL. Individuals on the SSL were 50% more likely to have at least one contact card and 39% more likely to have any interaction (including arrest, contact cards, victimization, court appearances, etc.) with the Chicago PD than their matched comparisons in the year following the intervention.”)

⁷⁷ *Id.* at 367.

⁷⁸ *Id.* at 363.

⁷⁹ Department of Procurement Services Non-Competitive Review Board (NCRM) Application: University of Chicago Crime Lab (Jan. 13, 2017), <https://www.chicago.gov/content/dam/city/depts/dps/SoleSource/NCRB2017/ApprovedNCRBUOCrimelab0217.pdf>.

⁸⁰ Allen Johnson Jr., *What the Studies Said*, NEW ORLEANS MAG. (May 2011), <http://www.myneworleans.com/New-Orleans-Magazine/May-2011/WHAT-THE-STUDIES-SAID/>.

⁸¹ Letter from Mitchell J. Landrieu, Mayor, City of New Orleans, to Eric H. Holder, Jr., Attorney Gen., U.S. Dep’t of Justice (May 5, 2010), http://media.nola.com/crime_impact/other/LettertoAttyGenHolder.050510.pdf.

⁸² U.S. DEP’T OF JUSTICE, CIVIL RIGHTS DIV., INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, vi-xiii (2011), https://www.justice.gov/sites/default/files/crt/legacy/2011/03/17/nopd_report.pdf [hereinafter INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT].

In 2013, the City of New Orleans and the Department of Justice entered a consent decree requiring structural and systemic reform with an independent monitor producing annual, quarterly, and special reports documenting compliance. These reports have mostly indicated good-faith, yet incremental progress, but the most recent Annual Report of the Consent Decree Monitor found the NOPD in non-compliance regarding stop, search and arrest practices.⁸³

The extreme scope of the NOPD's unlawful and biased practices between 2005 and 2011 is enough to cast doubt on all police and crime data, relevant to predictive policing systems, created during this period. In fact, the Department of Justice's Investigation Report, which substantiated the duration and scope of unlawful and biased practices, also documented evidence of "dirty data." The report identified several concerning disparities and inconsistencies in the NOPD arrest and Field Interview Cards data (which document NOPD encounters with residents, even those that do not result in arrest), and it questioned NOPD policies that encouraged unwarranted and potentially privacy-violating data collection.⁸⁴ For example, the report noted that in 2009, when NOPD arrest data was compared to national averages, "[t]he level of disparity for youth in New Orleans is so severe and so divergent from nationally reported data that it cannot plausibly be attributed entirely to the underlying rates at which these youth commit crimes, and unquestionably warrants a searching review and a meaningful response from the Department."⁸⁵ Additionally, the Department of Justice expressed concerns regarding omissions of essential information noting that the NOPD "[p]olicies and practices for complaint intake do not ensure that complaints are complete and accurate, systematically exclude investigation of certain types of misconduct, and fail to track allegations of discriminatory policing."⁸⁶

In 2012, the City of New Orleans entered a pro bono contract with the data-mining firm Palantir to use its proprietary Gotham data analysis and profiling services for crime-forecasting and to inform public safety strategies deployed by NOPD and other public safety agencies.⁸⁷ There is limited information about the Palantir system and its partnership with City of New Orleans because the contract and its subsequent extensions were entered without knowledge of key government officials and the public.⁸⁸ In fact, the lack of transparency about the use of the Palantir system has been the subject of a successful *Brady* challenge, citing the nondisclosure of the

⁸³ OFFICE OF THE CONSENT DECREE MONITOR, NEW ORLEANS, LA., 2017 ANNUAL REPORT OF THE CONSENT DECREE MONITOR FOR THE NEW ORLEANS POLICE DEPARTMENT CONSENT DECREE 9–11 (2018), <http://nopdconsent.azurewebsites.net/Media/Default/Documents/Reports/550%20MONITORS%20ANNUAL%20REPORT%20OF%202017.pdf>.

⁸⁴ INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, *supra* note 84, at 29–30.

⁸⁵ *Id.* at x.

⁸⁶ *Id.* at xvii.

⁸⁷ Presentation slides of two New Orleans government officials indicate that the Palantir system was used by NOPD to perform social network analysis and identify gang involvement in murders and shootings. It was also used by the Mayor's Office to locate neighborhood cleanup sites, and the Fire Department to increase its presence around schools in violent areas. SARAH SCHIRMER, OFFICE OF MAYOR MITCH LANDRIEU, DEPLOYING PALANTIR GOTHAM IN NEW ORLEANS, 11 (2014), <https://assets.documentcloud.org/documents/4344815/Nola-hc3-Final-20140403.pdf>; *see also* Ali Winston, *Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology*, VERGE (Feb. 27, 2018, 3:25 PM), <https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd>; PALANTIR TECHS., NOLA MURDER REDUCTION: TECHNOLOGY TO POWER DATA-DRIVEN PUBLIC HEALTH STRATEGIES (2014), <https://www.documentcloud.org/documents/4344816-NOLA-Murder-Reduction-White-Paper.html> (describing the Palantir and City of New Orleans partnership to identify "individuals exhibiting the highest predictors of violence").

⁸⁸ Winston, *supra* note 89.

system's analysis about the defendant and related gang activities.⁸⁹ Yet, presentation materials of two New Orleans government officials indicate that the Palantir system relied on NOPD data, including calls for service, electronic police reports, field information cards, and crime lab analysis, as well as data gleaned from the City's criminal⁹⁰ and non-law enforcement data sources, and open data sources, such as the location of liquor stores.⁹¹ There is no indication from available government and vendor documents that the NOPD data used to implement the system was scrubbed for errors and irregularities or otherwise amended in light of the dirty data identified in the Department of Justice report. In fact, evidence suggests that the Palantir system relied on some form of NOPD's dirty data because the system's analysis reflected similar racial disparities and other biases of the NOPD's practices and policies. City government documents highlighted that the NOPD system identified victims and perpetrators of violent or gang crimes as "overwhelmingly young, African American, male, undereducated, and underemployed," the same population that was disproportionately targeted by NOPD practices and overwhelmingly misrepresented in NOPD data.⁹² Though there can be additional or alternative explanations for this correlation, the scope and severity of NOPD's unlawful and biased practices and the extreme distortions identified in NOPD data suggests some level of attribution. The City of New Orleans has since cancelled its contract with Palantir in 2018, after public backlash regarding the secretive nature of the agreement.⁹³

D. Case Study 3: Maricopa County

In 2007, the ACLU filed a class-action lawsuit against the Maricopa County Sheriff's Office (MCSO) for alleged racially biased and unlawful police practices and policies as part of unlawful immigration enforcement operations,⁹⁴ which implicated contentious and unresolved legal questions regarding the authority and role of state and local police to enforce federal

⁸⁹ Defendant Kentrell Hickerson's Motion for Leave to Supplement His *Second Motion for New Trial* with Additional Grounds, Louisiana v. Hickerson, No. 516-272 (Mar. 8, 2018), <https://www.documentcloud.org/documents/4411697-Hickerson-Appeal-Defendant-s-Motion-to.html>; Matt Sledge, *Convicted Gang Leader Can Challenge NOPD's Use of Crime-Fighting Software, Judge Rules*, New Orleans Advocate (Mar. 14, 2018), https://www.theadvocate.com/new_orleans/news/courts/article_3a68a838-27bb-11e8-8b07-178e270926d4.html.

⁹⁰ Criminal data sources include, for example, calls for service, group and gang database, Sheriff's Office arrest and booking records, and probation and parole data.

⁹¹ The presentation slides of two New Orleans government officials indicate that the Palantir system was used to support criminal investigation, strategic homicide-reduction strategies, and to obtain indictments by employing multiple data sources. The presentation lists data sources including but limited to jail calls and phone data; gang affiliations and violent activity, crime lab data, social media. SCHIRMER, *supra* note 89; *see also* PALANTIR TECHS., *supra* note 89.

⁹² SCHIRMER, *supra* note 89.

⁹³ Ali Winston, *New Orleans Ends Its Palantir Predictive Policing Program*, VERGE (Mar. 15, 2018, 3:50 PM), <https://www.theverge.com/2018/3/15/17126174/new-orleans-palantir-predictive-policing-program-end>.

⁹⁴ Around January 2007, MCSO entered Memorandum of Agreement with U.S. Immigrations and Customs Enforcement (ICE) that purports to authorize enforcement of federal immigration laws by specially nominated and cross-trained MCSO staff. ICE stipulated that the agreement did not authorize MCSO staff to perform any of the biased and unlawful practices alleged in the ACLU complaint including, random street operations targeting day laborers and using race or immigration status as pretext for unlawful traffic stops. First Amended Complaint, Ortega Melendres v. Arpaio, No. CV 07-02513-PHX-MHM (D. Ariz. July 16, 2008), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-first-amended-complaint>.

immigration laws.⁹⁵ The following year, the Department of Justice announced an investigation of MCSO, but the investigation was delayed because MCSO refused to provide access to pertinent material and personnel.⁹⁶ In 2011, the Department of Justice released an investigation findings letter documenting MCSO's pattern of discriminatory behavior between 2007 and 2011, including discriminatory policing against Latino residents; unlawful stops and arrests; and unlawful retaliation against people who make complaints or criticize the department.⁹⁷ The Department of Justice also noted concerns that MCSO practices created "a wall a of distrust" that "substantially compromised effective policing by limiting the willingness of witnesses and victims to report crimes and speak to the police about criminal activity," which affects crime data and public safety within the County.⁹⁸ In 2013, a federal court found that MCSO engaged in unconstitutional and racially biased traffic stops and detention of Latino drivers, and it enjoined the MCSO from enforcing policies that permitted unlawful immigration enforcement.⁹⁹ In addition to detailing the scope of MCSO's biased practices, the decision noted significant irregularities and omissions in MCSO's records as well as evidence that MCSO officers and leadership openly, and often publicly, acknowledge biased and derogatory views and motives against Latino residents.¹⁰⁰ In response, a federal judge issued a court order mandating an annual review of MCSO practices and data, in addition to requiring more specific reforms.¹⁰¹

In 2015, the Department of Justice entered a settlement agreement¹⁰² with MCSO addressing some of the unlawful practices identified in its 2011 investigation, and joined the ongoing ACLU lawsuit as a plaintiff. In 2016, a federal court found several MCSO officers in civil contempt for deliberately violating the 2013 court order and continuing to engage in unconstitutional and discriminatory practices.¹⁰³ In compliance with the 2013 federal court order, MCSO commissioned Arizona State University (ASU) to perform annual reviews of its data. Unsurprisingly these reviews confirmed that this police data reflected the department's unlawful and racially biased practices. The two existing ASU annual reports of MCSO data covering 2014 to 2017 revealed that even while under consent decree the MCSO continued to engage in racially

⁹⁵ In its 2012 decision in *Arizona v. United States*, the Supreme Court held states are preempted from arresting or detaining individuals on the basis of suspected removability under federal immigration law. *Arizona v. United States*, 132 S. Ct. 2492 (2012). However, this decision did not address the legality of immigration inquiries that arise during the normal course of police activities unrelated to immigration enforcement, or the limitations on data generated by these legally questionable police activities. The proliferation of Sanctuary City laws and President Trump's aggressive and legally questionable immigration policies has further complicated these questions, and they currently remain unresolved.

⁹⁶ Letter from Thomas E. Perez, Assistant Att'y Gen., to Bill Montgomery, Maricopa Cty. Att'y, on the U.S.'s Investigation of the Maricopa County Sheriff's Office (Dec. 15, 2011), https://www.justice.gov/sites/default/files/crt/legacy/2011/12/15/mcso_findletter_12-15-11.pdf.

⁹⁷ *Id.*

⁹⁸ *Id.* at 16.

⁹⁹ Findings of Fact and Conclusions of Law, *Ortega Melendres v. Arpaio*, No. 2:07-cv-02513-GMS (D. Ariz. May 24, 2013), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-decision>.

¹⁰⁰ *Id.*

¹⁰¹ Supplemental Permanent Injunction/Judgment Order, *Ortega Melendres*, No. 2:07-cv-02513-GMS (D. Ariz. Oct. 2, 2013), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-order>.

¹⁰² Settlement Agreement, *United States v. Maricopa County*, No. 2:12-cv-00981-ROS (D. Ariz. July 17, 2015), <https://www.justice.gov/opa/file/631271/download>.

¹⁰³ Finding of Facts and Order Setting a Hearing for May 31, *Ortega Melendres*, No. 2:07-cv-02513-GMS (D. Ariz. May 13, 2016), <https://www.aclu.org/legal-document/ortega-melendres-et-al-v-arpaio-et-al-2016-order>.

biased and unlawful traffic stops and arrests of Black and Latino residents, with Latinos experiencing a greater likelihood of post-stop arrests or searches in 2015 to 2016.¹⁰⁴

While it is clear that Maricopa County's police data reflects its history of biased policing practices, it is also a case where it is difficult to know whether this data was used in a predictive policing system. There is no evidence of the MCSO using its own predictive policing system, but four cities within Maricopa County, which share data and police resources with MCSO,¹⁰⁵ are actively using predictive policing software, or have previously participated in a predictive policing pilot that may have relied MCSO data. In 2012, the Glendale Police Department in Maricopa County participated in a predictive policing pilot using RTMDx software, relying on Glendale police and crime data.¹⁰⁶ In this case, there is no evidence to suggest that MCSO data was directly used in this pilot. In fact, the MCSO and the Glendale Police Department only started officially sharing police and probation data in 2016, through AZ Link, a regional police data sharing platform which includes MCSO data dating from the period where the office is shown to have engaged in unlawful and biased practices and policies.¹⁰⁷ In 2014, the Tempe Police Department, in Maricopa County, received a Department of Justice Bureau of Justice Assistance grant to create a person and place based predictive policing pilot using several vendors software and services.¹⁰⁸ Though there is a study assessing the efficacy of this pilot, it does not indicate when the pilot occurred or the data the system used.¹⁰⁹ The Peoria Police Department in Maricopa County has used HunchLab predictive policing software since 2015, but there are no publicly available documents detailing the data the software is using or the City's data sharing policies with MCSO.¹¹⁰ Finally, in 2016, the Mesa Police Department in Maricopa County entered a three-year contract with the predictive policing software company, PredPol, which required the police department to provide local crime data.¹¹¹ A 2011 Mesa City Council document reveals that the Mesa Police Department uses AZ

¹⁰⁴ DANIELLE WALLACE ET AL., CTR. FOR VIOLENCE PREVENTION & CMTY. SAFETY, ARIZ. STATE UNIV., ANNUAL REPORT FOR THE MARICOPA COUNTY SHERIFF'S OFFICE: YEARS 2016 TO 2017, 25-31 (2018), https://cvpcs.asu.edu/sites/default/files/content/projects/MCSOreport2016_2017_0.pdf; DANIELLE WALLACE ET AL., CTR. FOR VIOLENCE PREVENTION & CMTY. SAFETY, ARIZ. STATE UNIV., PRELIMINARY YEARLY REPORT FOR THE MARICOPA COUNTY SHERIFF'S OFFICE, YEARS 2014-2015 (2016), <https://www.courthousenews.com/wp-content/uploads/2017/04/Maricopa-Bias-ASU-Report.pdf>.

¹⁰⁵ Many communities within Maricopa County rely on MCSO as its primary or backup law enforcement agency. The nature of this role is not articulated in available public documents, but Mesa, Tempe, parts of Peoria, and parts of Glendale are all listed as communities relying entirely or in part on MCSO services. *Districts*, MARICOPA COUNTY SHERIFF'S OFF., <https://www.mcso.org/Patrol/Districts#first> (last visited Feb. 1, 2019).

¹⁰⁶ LESLIE KENNEDY, JOEL CAPLAN & ERIC PIZA, CONJUNCTIVE ANALYSIS REPORT: 2012 ROBBERY IN GLENDALE, AZ (2015), http://www.rutgerscps.org/uploads/2/7/3/7/27370595/gpd_conjanalysis.pdf.

¹⁰⁷ See Memorandum of Understanding between Maricopa County Sheriff's Office and Glendale Police Department (Sept. 2016), <https://glendale-az.legistar.com/LegislationDetail.aspx?ID=2842290&GUID=9619C96A-1418-4455-A604-B85C399BE6AE&Options=&Search=>.

¹⁰⁸ BUREAU OF JUSTICE ASSISTANCE, U.S. DEP'T OF JUSTICE, CASE STUDY: TEMPE, ARIZONA, POLICE DEPARTMENT (2017), <https://it.ojp.gov/CAT/Documents/CaseStudyTempeArizonaPoliceDepartment.pdf>.

¹⁰⁹ See *id.*

¹¹⁰ David Baker, *Peoria Police Using Predictive Policing Technology to Stop Crime*, PEORIA TIMES (Jan. 2016), <http://raycomgroup.worldnow.com/story/30979418/peoria-police-using-predictive-policing-technology-to-stop-crime>.

¹¹¹ See Maria Polletta, *Can New Mesa Police Tool Prevent Crime from Happening?* (Oct. 21, 2016), <https://www.azcentral.com/story/news/local/mesa/2016/10/21/mesa-police-tool-prevent-crime-happening/89231252/>; City Council Report, Three-Year Term Contract for Predictive Policing Subscription for the Police Department (Aug. 22, 2016),

Link.¹¹² While it is evident that the Mesa Police Department uses MCSO data and that MCSO likely generates local police data as the backup law enforcement agency, it remains unconfirmed whether the department included this data directly in the crime data it provided to PredPol. It is difficult to make definitive conclusions regarding the use of MCSO's dirty data in these predictive policing systems because of the lack of publicly available information on the implementation of these systems in each jurisdiction and uncertainty regarding the role and relevant policies governing MCSO as a primary or backup law enforcement agency in each of these cities. Yet, this case study does highlight important concerns about the extraterritorial nature of police data, particularly when the practices and policies of relevant police departments are ill-defined and implicate controversial legal questions regarding the authority of local police.

III

THE PROBLEMS OF USING OF POLICE DATA AS CURRENTLY CONSTITUTED

A. Confirmation Feedback Loop

As the above examples show, numerous jurisdictions suffer under ongoing and pervasive police practices replete with unlawful, unethical, and biased conduct. This conduct does not just influence the data used to build and maintain predictive systems, it supports a wider culture of suspect police practices and ongoing data manipulation. Add to this the lack of oversight and accountability measures regarding police data collection, analysis, and use, and it becomes clear that any predictive policing system trained on or actively using data from jurisdictions with proven problematic conduct cannot be relied on to produce valid results without extensive independent auditing or other accountability measures. Yet police technology vendors have shown no evidence of providing this accountability and oversight, and other governmental actors rarely have the tools to do so. Thus, in such jurisdictions, these systems should be met with considerable suspicion that they are neutral, unbiased, or without risks of discrimination, and they should not replace or otherwise circumvent police reform measures.¹¹³

Though there is research that empirically demonstrates that the mathematical models of predictive policing systems are susceptible to runaway feedback loops, where police are repeatedly sent back to the same neighborhoods regardless of the actual crime rate, such feedback loops are also a byproduct of the biased police data.¹¹⁴ More specifically, police-data can be biased in two distinct ways. First and fundamentally, police data reflects police practices and policies. 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. Second, the data may omit essential information as a result of police practices and policies that

<https://mesa.legistar.com/LegislationDetail.aspx?ID=2808201&GUID=46F97E21-FFAB-41E1-846A-4F6C4B60F6E0&Options=&Search=>.

¹¹² CITY COUNCIL OF THE CITY OF MESA, AZ, COUNCIL MINUTES FOR AUGUST 29, 2011 (2011), <http://apps.mesaaz.gov/meetingarchive/archivedocuments/ClerkDetailedMinutes/August%2029,%202011%20RGL.pdf>.

¹¹³ See, e.g., DILLON REISMAN, JASON SCHULTZ, KATE CRAWFORD & MEREDITH WHITTAKER, AI NOW INST., ALGORITHMIC IMPACT ASSESSMENTS: A PRACTICAL FRAMEWORK FOR PUBLIC AGENCY ACCOUNTABILITY (Apr. 2018), <https://ainowinstitute.org/aiareport2018.pdf>.

¹¹⁴ Danielle Ensign et al., *Runaway Feedback Loops in Predictive Policing*, 81 PROC. OF MACHINE LEARNING RES. 1 (2018), <https://arxiv.org/abs/1706.09847>.

overlook certain types of crimes and certain types of criminals. For instance, police departments, and predictive policing systems, have traditionally focused on violent, street, property, and quality of life crimes. Meanwhile, white collar crimes are comparatively under-investigated and overlooked in crime reporting, despite a strong probability that they occur at a greater frequency than some of the other crime categories combined. Research on white collar crime is limited because the scope and nature of the crimes are constantly evolving so studies and surveys are often too narrow in focus, and complaints fail to reach or are not investigated by law enforcement. However, available studies estimate that approximately 49% of businesses and 25% of households have been victims of white collar crimes, compared to a 1.06% prevalence rate for violent crimes and 7.37% prevalence rate for property crime.¹¹⁵ Thus, while there is a significant need for more research on white collar crimes, available data demonstrates that these crimes occur at a greater frequency than crimes traditionally targeted by police departments and prominent predictive policing, like property and violent crimes.

The confluence of these distinct forms of skewed inputs ends up producing a questionable data-driven justification for increased policing and surveillance of historically overpoliced communities, and in turn reinforces popular misconceptions regarding the criminality and safety of underrepresented criminals and communities.¹¹⁶ The impact of these problems is most salient in the case study of New Orleans. The report on the Department of Justice investigation of the New Orleans Police Department revealed that officers improperly targeted and arrested transgender residents, sometimes fabricating evidence of a crime as well as exploiting archaic and biased statutes like “crimes against nature”—a statute that criminalizes sexual conduct that is considered morally unacceptable and requires registration as a sex offender.¹¹⁷ The Department of Justice found that in addition to these practices being discriminatory, they also raised significant concerns about the “efficient and effective use of resources to ensure public safety” since individuals convicted of “crimes against nature” made up approximately forty percent of the jurisdiction’s sex offender registry and the police department was charged with monitoring all registrants’ compliance.¹¹⁸ Moreover, since eighty percent of those registrants were also Black, the Department of Justice suggested there was an element of racial bias as well, which was confirmed by community members who told investigators “they believe some officers equate being African American and transgendered [sic] with being a prostitute.”¹¹⁹ The report also recognized the long-term consequences of the police department’s unlawful and biased practices noting “for the already vulnerable transgender community, inclusion on the sex offender registry

¹¹⁵ PwC, PULLING FRAUD OUT OF THE SHADOWS 5 (2018), <https://www.pwc.com/gx/en/forensics/global-economic-crime-and-fraud-survey-2018.pdf>; Gerald Cliff & April Wall-Parker, *Statistical Analysis of White-Collar Crime*, OXFORD RES. ENCYCLOPEDIA CRIMINOLOGY 7 (Apr. 2017), <http://oxfordre.com/criminology/view/10.1093/acrefore/9780190264079.001.0001/acrefore-9780190264079-e-267?print=pdf>; BUREAU OF JUSTICE STATISTICS, U.S. DEP’T OF JUSTICE, CRIMINAL VICTIMIZATION, 2016: REVISITED 11 (2018), <https://www.bjs.gov/content/pub/pdf/cv16re.pdf>; RODNEY HUFF ET AL., NAT’L WHITE COLLAR CRIME CTR., NATIONAL PUBLIC SURVEY ON WHITE COLLAR CRIME: 2010-14 (2010), https://www.nw3c.org/docs/research/2010-national-public-survey-on-white-collar-crime.pdf?sfvrsn=e51bbb5d_8.

¹¹⁶ See, e.g., Cynthia Lum, *The Influence of Places on Police Decision Pathways: From Call for Service to Arrest*, 28 JUST. Q. 631, 632 (2011) (“Place-based cues, especially those most noticeable to an officer (e.g. socioeconomic status, poverty, racial and ethnic makeup, disorder, crime, pedestrian and traffic density, and land use), may significantly affect an officer’s worldview and thereby his or her discretion.”).

¹¹⁷ INVESTIGATION OF THE NEW ORLEANS POLICE DEPARTMENT, *supra* note 84, at x.

¹¹⁸ *Id.*

¹¹⁹ *Id.*

further stigmatizes and marginalizes them, complicating efforts to secure jobs, housing, and obtain services at places like publicly-run emergency shelters.”¹²⁰

Confirmation feedback loops are so pernicious because they obfuscate the realities of crime and public safety. This is often magnified by public perceptions and public policy. When people observe increased police presence or contacts in marginalized communities it can reinforce unwarranted assumptions and stereotypes.¹²¹ Indeed, continued exposure or reinforcement of these stereotypes, especially in the absence of a counternarrative, can allow society to maintain a prejudice against marginalized groups while still maintaining an explicit commitment to egalitarianism.¹²² These complex, yet contradictory sentiments can incite responses that perpetuate this feedback loop. For instance, observing racially biased police practices can reinforce racial animus and false stereotypes of violence and criminality of certain racial or ethnic groups, which can result in improper calls for service for non-criminal activity that is perceived as suspicious or causes discomfort.¹²³ This is well documented in research on the social phenomena of “shopping while black”¹²⁴ in addition to the recent onslaught of media reports of white residents calling the police on black residents for non-criminal and innocuous actions like barbequing in a park or not smiling at a white neighbor.¹²⁵ This societal response of internalized bias and feedback loops is

¹²⁰ *Id.*

¹²¹ ANGELA J. HATTERY & EARL SMITH, *POLICING BLACK BODIES: HOW BLACK LIVES ARE SURVEILLED AND HOW TO WORK FOR CHANGE* 215–17, 232–34 (2017).

¹²² See JILLIAN OLINGER, KELLY CAPATOSTO & MARY ANNA MCKAY, *KIRWAN INST. FOR THE STUDY OF RACE & ETHNICITY, OHIO STATE UNIV., CHALLENGING RACE AS RISK* 21 (2017), <http://kirwaninstitute.osu.edu/wp-content/uploads/2017/02/implicit-bias-housing.pdf> (describing the causes and effects of implicit bias).

¹²³ See HATTERY & SMITH, *supra* note 123; Elijah Anderson, *The White Space*, 1 *SOC. RACE & ETHNICITY* 10, 13 (2015) (“In the absence of routine social contact between black and whites, stereotypes can rule perceptions, creating a situation that estranges blacks... In other words, whites and others often stigmatize anonymous black persons by associating them with putative danger, crime, and poverty...”); Elise C. Boddie, *Racial Territoriality*, 58 *UCLA L. REV.* 401, 434–46 (2010) (discussing how racial territoriality is a product of conscious and implicit biases that instigate territorial behavior that seeks to exclude people of color based of how they are perceived or represented); Angela Onwuachi-Willig, *Policing the Boundaries of Whiteness: The Tragedy of Being “Out of Place” from Emmett Till to Trayvon Martin*, 102 *IOWA L. REV.* 1113, 1170–83 (2017) (describing how White residents of the neighborhood where Trayvon Martin was murdered perceived Black people as intruders which resulted in repeated 911 calls despite the lack of evidence to support such suspicions).

¹²⁴ See, e.g., Shaun L. Gabbidon, *Racial Profiling by Store Clerks and Security Personnel in Retail Establishments*, 19 *J. CONTEMP. CRIM. JUST.* 345 (2003); Cassi Pittman, “*Shopping While Black*”: *Black Consumers’ Management of Racial Stigma and Racial Profiling in Retail Settings*, *J. CONSUMER CULTURE* (July 27, 2017), <https://journals.sagepub.com/doi/abs/10.1177/1469540517717777>; George E. Schreer, Sandra Smith & Kirsten Thomas, “*Shopping While Black*”: *Examining Racial Discrimination in a Retail Setting*, 39 *J. APPLIED SOC. PSYCHOL.* 1432 (2009).

¹²⁵ See, e.g., Carla Herreria, *Woman Calls Police on Black Family for BBQing at a Lake*, *HUFFINGTON POST* (May 11, 2018, 10:24 AM), https://www.huffingtonpost.com/entry/woman-calls-police-oakland-barbecue_us_5af50125e4b00d7e4c18f741 (describing how a black family was detained and questioned by police after a white woman called the police on them for attempting to barbeque at a public lake in Oakland, California); Amber Jamieson, *A White Woman Called the Police on Bob Marley’s Granddaughter for Not Smiling at Her*, *BUZZFEED NEWS* (May 10, 2018, 3:04 PM), <https://www.buzzfeednews.com/article/amberjamieson/black-artists-airbnb-white-woman-police-cops#.rwYRXPYOg> (describing how three black people were detained by police after a white resident called the police because they failed to smile, wave back, and acknowledge her); P.R. Lockhart, *A Black Political Candidate Was Canvassing in her District. Then Someone Called the Police.*, *VOX* (Sept. 20, 2018, 3:20 PM), <https://www.vox.com/identities/2018/9/20/17883018/shelia-stubbs-canvassing-911-police-racial-profiling-wisconsin> (describing how police received a false complaint of suspected drug activity when a black state legislator was canvassing in her district).

also becoming more prevalent as historically segregated and majority non-white and lower income neighborhoods experience gentrification. There is a growing body of evidence documenting heightened neighbor-driven police enforcement in gentrifying neighborhoods.¹²⁶ Most recently, a study on 311 calls for service data in New York City found that lower-income communities of color with the largest influxes of white residents experienced significantly higher increases in quality of life complaints, and summons and arrests outcomes were three times more likely than in neighborhoods without large influxes of white residents.¹²⁷

These societal responses to the feedback loop contribute to policing's dirty data problem in several distinct ways, but the most concerning influence, especially in the predictive policing context, is that calls for service provides several opportunities for discretion and selective enforcement, which can further distort police data. There is not only a great amount of subjectivity permitted in assessing the validity of a calls for service and whether to involve law enforcement, but once police are present, they have full discretion in how to respond (e.g. whether to arrest individuals involved in the incident or de-escalate) and how to report their interactions.¹²⁸ This includes whether to classify their interactions as crimes, infractions, or a non-criminal incident. All of these subjective decisions alter police data that is often used in predictive policing systems (e.g. crime and calls for service data), and both longstanding research on police officer discretion and more recent research on calls for service outcomes demonstrate that police bias and institutional interests are reflected in the outcomes and reporting, despite actual crime levels and neighborhood conditions.¹²⁹ Thus, when the dirty data generated by dubious calls for services and

¹²⁶ See, e.g., Ayobami Laniyonu, *Coffee Shops and Street Stops: Policing Practices in Gentrifying Neighborhoods*, 54 URB. AFF. REV. 898 (2017) (presenting empirical analysis that demonstrates a strong positive association between gentrification and the adoption of punitive policing strategies); Lam Thuy Vo, *They Played Dominoes Outside Their Apartment For Decades. Then The White People Moved In And Police Started Showing Up.*, BUZZFEED NEWS (June 29, 2018), <https://www.buzzfeednews.com/article/lamvo/gentrification-complaints-311-new-york#.xoJoxzbEA>; *311 Reports in SF by Neighborhood 2008-16*, Anti-Eviction Mapping Project, <http://www.antievictionmappingproject.net/311.html> (last visited Feb. 4, 2019) (noting a disproportionate increase in 311 calls in gentrifying neighborhoods in San Francisco).

¹²⁷ *New Neighbors and the Over-Policing of Communities of Color*, CMTY. SERV. SOC'Y (Jan. 6, 2019), <http://www.cssny.org/news/entry/New-Neighbors>.

¹²⁸ Lum, *supra* note 118, at 640, 643 (illustrating police officer decision pathways for calls for service).

¹²⁹ See, e.g., David A. Klinger & George S. Bridges, *Measurement Error in Calls-for-Service as an Indicator of Crime*, 35 *Criminology* 705 (1997) (concluding that estimates of neighborhood crime rates based on calls for service and police dispatch data is biased and can be misleading because the data is embedded with errors and subject to police undercounting crime based on neighborhood or community characteristics); Lum, *supra* note 118, at 635 ("Places with high proportions of Black residents may be treated differently by officers than places that are predominately White, even if both places have similar levels of social disorganization or crime."); ERIC J. SCOTT, NAT'L INST. OF JUSTICE, *CALLS FOR SERVICE: CITIZEN DEMAND AND INITIAL POLICE RESPONSE* (1981), <https://www.ncjrs.gov/pdffiles1/Digitization/78362NCJRS.pdf> (finding non-white individuals and women are more likely to be arrested under circumstances that are found not to constitute sufficient grounds for prosecution); Douglas A. Smith, Christy A. Visher & Laura A. Davidson, *Equity and Discretionary Justice: The Influence of Race on Police Arrest Decisions*, 75 *J. CRIM. L. & CRIMINOLOGY* 234 (1984) (finding that in police encounters where officers have discretion Black people were more likely to be arrested than White people, after controlling for other factors); Lam Thuy Vo, *They Played Dominoes Outside Their Apartment for Decades. Then the White People Moved in and Police Started Showing Up.*, *BuzzFeed News* (June 29, 2018), <https://www.buzzfeednews.com/article/lamvo/gentrification-complaints-311-new-york#.xoJoxzbEA> (finding police responses to 311 calls increased sixfold as a Harlem neighborhood experienced increased gentrification).

subjective police enforcement and reporting is used in predictive policing systems, the technology can produce predictions that further perpetuate confirmation feedback loops.

Confirmation feedback loops also influence public policy by driving or providing justification for government officials to support policies that attempt to micromanage or push out communities that are misperceived as producing problems or increasing disorder. This phenomena of artificially manufactured moral panic and public consent of new forms of state control was interrogated in the groundbreaking book, *Policing the Crisis*. Stuart Hall and his co-authors conducted an empirical study of the social construction of street crime, and the societal labeling of the black men as “muggers.” The authors found that despite public belief that there was a new street crime pandemic, there was little evidence to support this belief. Instead, they argued, street crime was not new but manufactured as a new problem by media, which then influenced which people police identified as criminals, and reinforced biases of judges who created justifications for the state-sponsored control of the black community.

Now this phenomenon is commonly experienced though austerity measures and policies that criminalize the conditions that contribute to the marginalization of some communities over others.¹³⁰ Common examples are nuisance laws and ordinances, which empower municipal governments to penalize individuals and communities for a certain number of calls for service or alleged “nuisance” conduct, which is an ill-defined category of conduct that can range from assault to littering depending on the jurisdiction.¹³¹ A recent New York Civil Liberties Union report found that these policies disproportionately affect poor communities of color because they “amplify the harms of the criminal justice system and exacerbate socioeconomic and racial inequalities by making housing instability a consequence of law enforcement.”¹³² These policies and practices result in the displacement of these communities and are often surreptitiously pursued in order to attract private investment and consumption.¹³³

B. The Boundless Reach of Biased Police Data

These case studies demonstrate that without an empowered and independent authority, the potentially unlawful and biased practices and policies of police departments as well as the subsequent data produced through these practices can remain unaddressed and uncorrected. When dirty data is fed into a new predictive system, it can fundamentally taint its recommendations. This can further ingrain biases in supposedly “neutral” systems. This is important since there are few political and institutional incentives that encourage self-monitoring and reform, or ongoing

¹³⁰ Casey Kellogg, *There Goes the Neighborhood: Exposing the Relationship Between Gentrification and Incarceration*, 3 THEMIS: RES. J. JUST. STUD. & FORENSIC SCI. 174, 184 (2015) (“As migration patterns lead to more homogenous communities, it becomes easier to label the behavior of racial and class groups as criminal.”).

¹³¹ See Anna Kastner, Note, *The Other War at Home: Chronic Nuisance Laws and the Revictimization of Survivors of Domestic Violence*, 103 CALIF. L. REV. 1047, 1052 (2015) (“Nuisance law is a well established, but vague, common law doctrine. According to Prosser, ‘[t]here is general agreement that it is incapable of any exact or comprehensive definition,’ because it contains a mixture of criminal law, tort law, and property law.”); N.Y. CIVIL LIBERTIES UNION, MORE THAN A NUISANCE: THE OUTSIZED CONSEQUENCES OF NEW YORK’S NUISANCE ORDINANCES 6 (2018), https://www.nyclu.org/sites/default/files/field_documents/nyclu_nuisancereport_20180809.pdf [hereinafter MORE THAN A NUISANCE] (describing the range of conduct that can be included in a nuisance ordinance).

¹³² MORE THAN A NUISANCE, *supra* note 133, at 10.

¹³³ Donald C. Bryant Jr. & Henry W. McGee Jr., *Gentrification and the Law: Combatting Urban Displacement*, 25 WASH. U. J. URB. & CONTEMP. L. 43, 70–72 (1983); Jefferson, *supra* note 20, at 3–4; Kellogg, *supra* note 132, at 192–95.

auditing of data systems.¹³⁴ We can see this in the failure of NOPD to self-reform after the first Department of Justice investigation into the department’s practices, and similarly in the MCSO’s persistent defiance of the federal court order.

These case studies also demonstrate that merely identifying unlawful and biased practices is not enough. The data collection, analysis, and use practices must be reformed as well. Though most of the jurisdictions reviewed in our research engaged in some level of data collection and review reforms, none of the legal agreements restricted the use of the data generated during the periods of unlawful and biased police practices, which would be a meaningful limitation in future legal agreements on police reform. Thus, restrictions or prohibitions on the use of the historical data generated by unlawful and bias practices are necessary to ensure that the legacy of such practices is not perpetuated through the systems that rely on such data.

Police data generated by the unlawful or biased practices and policies of a specific police department or division can corrupt practices and data in other jurisdictions, and skew decision-making throughout the criminal justice system, often in ways that are difficult to account for and correct. The Maricopa County case study demonstrates additional risks from dirty data production – the risks that other jurisdictions, including ones that have not been found to systematically violate the law, may incorporate such data into their own predictive policing systems, potentially corrupting additional new data and practices. Data sharing between police departments occurs frequently and in some cases is encouraged and given federal government support.¹³⁵ And data sharing is not limited to law enforcement agencies. Police and crime data are used in decision-making during prosecution, pretrial services, adjudication, sentencing, and corrections, as well as non-criminal justice related political decisions, such as community investment and development. The lack of transparency and oversight regarding police practices, policies, and the data created through these raises serious concerns regarding the possibility that *any* predictive system relying on police data could operate in a fair and just manner.

IV CONCLUSION

Data is seen as an important tool for policy making and governance because in its absence there is often too much reliance on subjective factors. The last twenty years have seen a widespread adoption of data-driven practices, policies, and technologies in the public sector. Yet this increasing reliance on data to assess and make decisions about complicated social, economic, and

¹³⁴ See, e.g., Friedman & Ponomarenko, *supra* note 23 (highlighting why police are immune from regulation or constraints by other branches of government and other unique challenges in regulating police); Myriam E. Gilles, *Breaking the Code of Silence: Rediscovering “Custom” in Section 1983 Municipal Liability*, 80 B.U. L. REV. 17, 31 (2000) (“Aside from the overdeterrence of individual officers, it seems clear that where liability falls solely on individual officers, municipalities have little incentive to develop comprehensive responses to rampant unconstitutional practices.”); Stephen Rushin, *Using Data to Reduce Police Violence*, 57 B.C. L. REV. 117, 152–54 (describing how municipal unwillingness to allocate resources for police reform, collective bargaining, and civil service protections inadvertently discourage police leadership from proactively and forcefully responding to misconduct).

¹³⁵ RACHEL LEVINSON-WALDMAN, BRENNAN CTR. FOR JUSTICE, WHAT THE GOVERNMENT DOES WITH AMERICANS’ DATA 13-15 (2013), <https://www.brennancenter.org/publication/what-government-does-americans-data>; *About the RISS Program*, REGIONAL INFORMATION SHARING SYSTEM, <https://www.riss.net/about-us/> (last visited Feb. 1, 2019).

political issues presents serious risks to fairness, equity, and justice, if greater scrutiny is not given to the practices underlying the creation, auditing and maintenance of data.

Our research demonstrates the risks and consequences associated with overreliance on unaccountable and potentially biased data to address sensitive issues like public safety. These case studies show that illegal police practices significantly distort the data that is collected, but that data is still used for law enforcement and other purposes. The failure to adequately interrogate and reform police data creation and collection practices can result in skewed predictive policing systems and create lasting consequences that will permeate throughout the criminal justice system and society more widely.

There may be a natural inclination to assume that predictive policing vendors can address the problems of dirty data identified in this study by removing known cases. But such mitigation methods are inadequate for a number of reasons. First, since there are few incentives and almost no requirements for police departments to self-monitor and reform practices or policies that create biased or dirty data, it is unlikely that police departments would identify these problems for a vendor to remove or otherwise address. Second, there is no current methodology or mechanisms for identifying these problematic practices and policies in real-time; therefore, it is impossible to distinguish substantiated and suspected problems. Third, as we have argued, a fundamental flaw of police data is that it does not capture all relevant crime information because of institutional policies or practices that ignore certain types of crimes or criminals, negative community relations that affect which crimes the police track, and corrupt or unethical practices that lead to the omission or manipulation of police records. There is no meaningful way for a vendor to adjust what is unknown or not recorded. The absence of data is as significant as its creation: yet there is no technical “fix” for this. Instead, mitigation efforts should be focused on developing reliable mechanism for assessing the harms inherent in the use of historical police data, as well as data generated after implementation of police data collection reform, and backed by strong public transparency and accountability measures.

The jurisdictions researched for this paper were limited to police departments that were subjects of publicized investigations and federal litigation. These case studies demonstrate the importance of independent government investigations and federal court litigation in uncovering unlawful and biased police practices that would otherwise persist without federal government intervention. Yet, these crucial mechanisms for uncovering and addressing problematic police practices have been threatened with the parting acts of former U.S. Attorney General Jeff Sessions just before his unexpected forced resignation.¹³⁶ Before he left office, he issued a Department of Justice policy memo significantly limiting the use of consent decrees by requiring top political appointees to sign off, limiting their scope and duration, and requiring department attorneys to provide evidence of additional violations *beyond* unconstitutional behavior.¹³⁷ These limitations are significant and serious. The result may be that problematic police departments will remain unchecked, not because of lack of evidence of unconstitutional practices, but because the new standard of evidence is extremely high or because political leadership refuses to sign off. In light of these developments and the absence of incentives for self-scrutiny and reform, collective action for greater accountability, oversight, and redress is urgent. A broad coalition of stakeholders is needed to push public discourse on the drivers and consequences of dirty data, and to motivate

¹³⁶ See Katie Benner, *Sessions, in Last-Minute Act, Sharply Limits Use of Consent Decrees to Curb Police Abuses*, N.Y. TIMES (Nov. 8, 2018), <https://www.nytimes.com/2018/11/08/us/politics/sessions-limits-consent-decrees.html>.

¹³⁷ *Id.*

government officials to act to ensure that principles of fairness, equity and justice are reflected in government practices.

APPENDIX

Jurisdiction	Type of Agreement (Year)	Unlawful/Bias Practices	Timeframe of Unlawful/Bias Practices	Predictive Policing System (Year)	Evidence of bias data in Predictive Policing
Baltimore (MD)	DOJ Consent Decree (2017)	Unconstitutional and racially biased stops, searches, and arrests; excessive use of force; retaliation against individuals for constitutionally-protected expression	2010–16	Baltimore City- In development; Baltimore County- several pilots	Unclear. Evidence suggest BPD and County Police shared data that could have been used in the County’s pilots. BPD plans to acquire or develop a predictive policing system.
Boston (MA)	Boston Police Department and ACLU Commissioned Investigation (2015)	Racially discriminatory stop and frisk practices	2007–10	NIJ-grant (2009 retrospective review of Safe Streets Teams); 2010 BPD official stated plans to develop a system	Safe Streets Teams system used 2000–09 BPD data.
Chicago (IL)	Settlement Agreement (2015), DOJ Memorandum of Agreement (2018), IL Attorney General Consent Decree (2019)	Pattern and practice of unconstitutional and racially biased stops, searches, and arrests; excessive use of force; poor data collection to identify and address unlawful conduct	2011–16	RTMDx (2012); Strategic Subjects Lists (2012-present)	Both systems used CPD arrests and other police data
Ferguson (MO)	DOJ Consent Decree (2016)	Pattern or practice of unlawful and discriminatory stops, searches, and arrests; excessive use of force, First amendment violations. DOJ noted similar problems in St. Louis County.	2007–16	HunchLab (St. Louis County, 2015–present)	Unclear. County police perform police work in and around Ferguson and there is overlap between unlawful practices and County’s predictive policing use. No public policies regarding data sharing between County and Ferguson
Miami (FL)	DOJ Memorandum of Agreement (2016)	Pattern and practice of excessive force with use of firearms. Prior to DOJ investigation, several MPD officers were indicted on conspiracy charges for lying and planting evidence to undermine investigations of officer-involved shootings	2008–11	HunchLab (2014–17), IBM (2011)	IBM system used historical crime data

Maricopa County (AZ)	Federal District Court Order (2013 and 2016), DOJ Consent Decree (2015)	Unconstitutional and racially biased stops, searches, and arrests; unconstitutional lengthening of stops; unlawful retaliation against people who made complaints or criticized MCSO	2007–11; 2014–17	PredPol (Mesa, 2016–Present); HunchLab (Peoria, 2015–Present); RTMDx (Glendale 2012 pilot); federally-funded pilot (Tempe, 2014)	Unclear.
Milwaukee (WI)	Settlement Agreement with five-year Consent Decree (2018)	Unconstitutional and racially biased stop and frisk practices. Draft DOJ investigation report found racial disparities in traffic enforcement practices	Since 2008	Milwaukee County received federal funds for data-driven policing	Prior police chief expressed interest in using predictive policing. Unclear if County predictive policing system uses MPD data.
New Orleans (LA)	DOJ Consent Decree (2013)	Pattern or practice of excessive use of force; unlawful stops, searches, and arrests; discrimination based on race, national origin, and LGBT status; gender discrimination in failure to investigate violence against women	2005–11	Palantir (2012–18)	Palantir used NOPD data to train and use system.
New York (NY)	Remedial Process (2013), Settlement Agreement (2015)	Unconstitutional and racially biased stops, searches and arrests	2003–12	HunchLab (2015 pilot); NIJ-funded pilot (funded in 2013 for three-year grant); Palantir (unknown)	HunchLab used historical NYPD crime data
Newark (NJ)	DOJ Consent Decree (2016, amended in 2018)	Pattern or practice of unlawful stops, searches and arrests; racially biased policing practices; excessive use of force; theft by NPD officers	2007–12	Starlight (unknown); PredPol (unknown); RTMDx (2012 Pilot)	Pilot occurred during DOJ investigation so it is possible system used biased NPD data. PredPol uses NPD historical crime location data
Philadelphia (PA)	Settlement Agreement and Consent Decree (2011)	Unconstitutional and racially biased stop and frisk practices; excessive use of force; deficient and inconsistent supervision and operational controls	2005–10, Jan.–June 2017	federal-funded pilot (2010); HunchLab (2013–Current)	PPD data used in pilot. Recent finding of bias practices can be reflected in data currently being used in Hunchlab
Seattle (WA)	DOJ Consent Decree (2012)	Pattern or practice of excessive use of force. DOJ did not issue a finding of	2009–11	PredPol (2013–Present)	SPD documents state the PredPol system uses crime data dating back to 2008

		racially biased policing but expressed serious concerns about bias particularly in pedestrian encounters			
Suffolk County (NY)	DOJ Memorandum of Agreement (2014)	Bias-based policing against Latinos; deficiencies in reporting of police activities	2004–11	Vendor unknown, but federal and philanthropic funds were awarded to develop or acquire a system	Media reports indicate use of a predictive policing system since 2012 but vendor unknown.