2023

# When Dirty Data Leads to Dirty Policing

Madison Blevins
*University of Richmond - School of Law*

Richmond Journal of Law & Technology Volume XXIX, Issue 2

# WHEN DIRTY DATA LEADS TO DIRTY POLICING

Madison Blevins*

Cite as: Madison Blevins, *When Dirty Data Leads to Dirty Policing*, 29 RICH. J.L. & TECH. 166 (2023)

* J.D. Candidate, University of Richmond School of Law, 2023. B.S., University of South Carolina, 2020. I would like to thank Professor Corinna Barrett Lain for her guidance and continued support in the writing of this article. She helped me find my interest in criminal procedure and constitutional law, and I am so grateful for her help throughout this learning process. I would also like to thank the amazing editors and staff of the Richmond Journal of Law & Technology for their efforts in bringing this article to publication.

## INTRODUCTION

[1] On May 25th, 2020, George Floyd was tragically killed by police officers in Minneapolis.[1] While George Floyd's death was the shock that catapulted the Black Lives Matter ("BLM") movement to the center of international attention,[2] it was also just the tip of the iceberg. Floyd's death was not the first death of a black person at the hands of the police, nor would it be the last. "A black person is killed by a police officer in America at a rate of more than one [person] every other day."[3] These repeated incidents across the country have ignited a mass movement centered on police violence against people of color, and predictive policing is at the forefront of the conversation.[4] Yet the timing and casual cruelty of the death of George Floyd, recorded and shared on social media, spurred a national uprising. As people across America protested in the streets, the public seemed to take a greater interest in the history of the American criminal justice system and its roots in racial oppression. Although BLM has existed since 2013, the movement and policy discussion has gained a great deal of attention since the summer of 2020.

---

[1] Alex Altman, *Why The Killing of George Floyd Sparked an American Uprising*, TIME (June 4, 2020, 6:49 AM), https://time.com/5847967/george-floyd-protests-trump/ [https://perma.cc/J3RU-ZBW4].

[2] George Floyd was not the only one, unfortunately. Other salient police killings include Breonna Taylor, Eric Garner, and Tamir Rice, among others. *See George Floyd: Timeline of black deaths and protests*, BBC (Apr. 22, 2021), https://www.bbc.com/news/world-us-canada-52905408 [https://perma.cc/CX2V-KYRZ].

[3] Altman, *supra* note 1.

[4] *See* Ram Subramanian & Leily Arzy, *State Policing Reforms Since George Floyd's Murder*, BRENNAN CTR. FOR JUST. (May 21, 2021), https://www.brennancenter.org/our-work/research-reports/state-policing-reforms-george-floyds-murder [https://perma.cc/LHY2-W4FH].

[2] Predictive policing has received public criticism for its problematic use of data and artificial intelligence ("AI") to predict where crime is most likely to occur and who is most likely to commit it.[5] While it may seem like artificial intelligence could have a positive impact on both implicit and explicit biases present in policing, an AI system is only as good as the data it uses. Welcome to the world of dirty data: data that is flawed in some way.[6]

[3] This paper will argue that predictive policing fed by dirty data has created an almost foolproof way to justify police suspicion of any citizen who finds themselves a target of predictive policing lists, whether by location or person-specific means. These systems not only produce skewed outcomes, but they also let the police write a blank check to do the very things that objective justifications, like probable cause and reasonable articulable suspicion, are designed to keep the police from doing. Predictive policing will arguably always give the police an articulable reason or cause for suspicion, eviscerating Fourth Amendment protections.

[4] Despite major problems in the world of artificial intelligence, the problem of dirty data has received relatively little attention from courts and commentators. Few defendants are making Fourth Amendment challenges on this basis, so courts are not ruling on it.[7] For their part, commentators are

[5] *See generally* Johana Bhuiyan, *LAPD ended predictive policing programs amid public outcry. A new effort shares many of their flaws*, THE GUARDIAN (Nov. 8, 2021, 1:00), https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform [https://perma.cc/A235-8XSL] (discussing the harms of predictive policing).

[6] Rashida Richardson et al., *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. REV. 192, 195 (2019).

[7] *See generally* Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109 (2017) (addressing the Fourth Amendment, Selbst's article focuses more on disparate impact of racial profiling and less on the impacts this article will discuss); Renata M.

not paying attention either.[8] Scholarship in this area has generally focused on the problematic use of algorithms in the context of sentencing, probation, and facial recognition—not the dirty data underlying it.[9] In short, existing scholarship has not seen the connection between the two sets of problems.

[5] This paper fills that gap by exploring the problems with dirty data as it populates artificial intelligence generally and predictive policing more specifically. It describes how the government uses dirty data in its predictive policing, why that is worrisome, and what the constitutional implications are of these arrangements. Part I addresses what dirty data and artificial intelligence are and how these two concepts are interconnected. Part II explores predictive policing to build a foundation to fully understand the problems artificial intelligence poses in the predictive policing context. Finally, Part III explores the constitutional implications of issues created by dirty data, examining the troubles dirty data presents in the Fourth Amendment context. To address the troubles of predictive policing, we must also address the problems of the dirty data that drive it.

## II. DIRTY DATA AND ARTIFICIAL INTELLIGENCE

[6] To understand the problems with predictive policing, we first must understand the dirty data and artificial intelligence behind the algorithms that support predictive policing. Part I supplies that foundation. Section A

O'Donnell, Note, *Challenging Racist Predictive Policing Algorithms Under the Equal Protection Clause*, 94 N.Y.U. L. REV. 544 (2019) (discussing Equal Protection Clause implications); Sarah Valentine, *Impoverished Algorithms: Misguided Governments, Flawed Technologies, and Social Control*, 46 FORDHAM URB. L.J. 364 (2019) (discussing how the equal protection claim has failed in previous cases).

[8] *See* Selbst, *supra* note 7, at 146; Valentine, *supra* note 7, at 408.

[9] *See* Selbst, *supra* note 7, at 113; O'Donnell, *supra* note 7, at 547; Valentine, *supra* note 7, 365–370.

provides the fundamentals on dirty data, and Section B provides the foundation of artificial intelligence.

### A. Dirty Data

[7] Dirty data is a term used within the data mining and research community to refer to "missing data, wrong data, and non-standard representations of the same data."[10] The term dirty data also "includes data that is derived from or influenced by corrupt, biased, and unlawful practices, including data that has been intentionally manipulated or … distorted by individual and societal biases."[11] Because data is subject to more than one form of input and manipulation simultaneously, it can be difficult for systems to detect and separate good data from bad data.[12] This is especially true when the data production process itself is biased or otherwise part of the problem.[13]

[8] A straightforward way to think of dirty data is GIGO – "garbage in, garbage out."[14] In computing and other data-related spheres, this phrase is used to express the idea that "incorrect or poor quality input[s] will . . . produce [a] faulty output."[15] If there are issues with the data being fed into

[10] Richardson et al., *supra* note 6, at 195 (citing Won Kim et al., *A Taxonomy of Dirty Data*, 7 DATA MINING & KNOWLEDGE DISCOVERY 81, 81 (2003)).

[11] *Id*.

[12] *Id*. at 196.

[13] *Id*.

[14] *overview: garbage in garbage out*, OXFORD REFERENCE, https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095842747 [https://perma.cc/YY7P-A4XV].

[15] *Id*.

a system, then there will be issues with the information being generated. This is particularly troubling in the criminal justice space, although it plagues all kinds of data.[16]

[9] Dirty data stems from routine human error and leads to four main weaknesses.[17] First, "[p]eople can make mistakes in data collection, input, [and] integration of datasets," which means the data being relied on is incorrect.[18] Second, data can be incomplete and contain missing fields or records, since the input of data largely relies on human actors.[19] Third, data can be inconsistent; it can "involv[e] overlapping codes or code meanings that change over time."[20] Fourth, data can be incomprehensible, containing formatting issues or the inclusion of multiple data points in a single field.[21] In order to overcome these weaknesses, data should be "scrubbed" or "cleaned" to allow for ethical use.[22]

[10] Despite how easy it is for humans to make these mistakes when inputting data, the associated risks presented by dirty data are neither abstract nor minute. The threat of dirty data extends beyond the risk that the data will not provide reliable information. Rather, individual liberties are

[16] *See* Vincent M. Southerland, *The Intersection of Race and Algorithmic Tools in the Criminal Legal System*, 80 MD. L. REV. 487, 505 (2021).

[17] *Id*.

[18] *Id*.

[19] *Id*.; Wayne N. Renke, *Who Controls the Past Now Controls the Future: Counter-Terrorism, Data Mining and Privacy*, 43 ALTA. L. REV. 779, 791 (2006).

[20] Renke, *supra* note 19, at 791.

[21] *Id*.

[22] *See id*. at 792.

threatened when data mining produces inaccurate and unreliable data.[23] Although some associated risks come from the potential misuse of technology, other risks and problems with dirty data are inherent in even the best-intentioned cases.[24] No matter how data is used, data mining generates social, political, and personal risks. Pattern-based data mining enables the government to have widespread access to an individuals' personal information.[25] This type of pervasive access to personal information allows the government to develop profiles and then run profiles against this information, thereby creating reasonable articulable suspicion.[26]

[11] With such widespread access to and reliance on data mining, it is easy to see how dirty data can have a detrimental effect on society in the criminal justice context. In fact, data errors leading to faulty predictions and potentially dangerous incorrect decisions have previously been addressed by the Supreme Court of the United States.[27] In 2009, Justice Ruth Bader Ginsburg warned that "[i]naccuracies in expansive, interconnected collections of electronic information raise grave concerns for individual liberty."[28] When dirty data is used to create artificial intelligence, these systems incorporate the corrupt data, allowing law enforcement to use it in a problematic way.[29] To understand how that happens, one must understand

[23] *See id*. at 795.

[24] *Id*. at 795.

[25] Renke, *supra* note 19, at 796.

[26] *See id*. at 796.

[27] *See* Valentine, *supra* note 7, at 389.

[28] *Id*. (citing Herring v. United States, 555 U.S. 135, 155 (2009) (Ginsburg, J., dissenting)).

[29] *See, e.g.*, *id*. at 390.

how dirty data feeds into artificial intelligence—the next topic of discussion.

### B. Dirty Data in Artificial Intelligence

[12] Artificial intelligence is the "science and engineering of making intelligent machines, especially intelligent computer programs."[30] Artificial intelligence is related to the "task of using computers to understand human intelligence, but . . . does not have to confine itself to methods that are biologically observable."[31] In its most simple form, "artificial intelligence is a field that combines computer science and robust datasets to enable problem-solving."[32] Artificial intelligence seeks to create systems that "make predictions or classifications based on input data."[33] Within artificial intelligence, there are two recognized approaches: the human approach and the ideal approach.[34] The human approach includes systems that think like humans and act like humans. The ideal approach includes systems that think and act "rationally," with less human-like emotions and mistakes.[35] Artificial intelligence systems are "powered by algorithms, using

[30] John McCarthy, *What is Artificial Intelegence?*, https://www-formal.stanford.edu/jmc/whatisai.pdf [https://perma.cc/99ZM-HJUW] (last revised Nov. 12, 2007).

[31] *Id*.

[32] IBM Cloud Education, *Artificial Intelligence (AI)*, IBM CLOUD (June 3, 2020), https://www.ibm.com/cloud/learn/what-is-artificial-intelligence [https://perma.cc/7AMJ-8SE7].

[33] *Id*.

[34] *Id*.

[35] *See id*.

techniques such as machine learning and deep learning to demonstrate 'intelligent' behavior."[36]

[13] Machine-learning is the process by which computers develop pattern recognition; the "ability to continuously learn from and make predictions based on data."[37] Machine-learning systems are also able to "adjust without being specifically programmed to do so."[38] Finally, machine-learning "automates the process of analytical model-building and [enables] machines to adapt to new scenarios" and situations, independent from human programming.[39]

[14] To build a machine-learning artificial intelligence system requires four steps.[40] The first step is to "select and prepare a training data set necessary to solving a problem" chosen by the system's designer.[41] The second step is to choose an algorithm for the training data.[42] Depending on whether the data is labeled or unlabeled, the algorithm could be a regression, decision tree, clustering algorithm, association algorithm, or a neural

[36] *What is Artificial Intelligence*, HEWLETT PACKARD ENTER., https://www.hpe.com/us/en/what-is/artificial-intelligence.html?jumpid=ps_8m6wvisfq7_aid-520061736&ef_id=Cj0KCQiA64GRBhCZARIsAHOLriIPKCWb4hE_7gyIqfPsO853BQ89gKYbWmqzAVX0IDmK3RliRFYOWhYaAtbREALw_wcB:G:s&s_kwcid=AL!13472!3!558204189304!e!!g!!what%27s%20artificial%20intelligence!13236197162!129170842076& [https://perma.cc/BP2B-HUDE] [*hereinafter AI*].

[37] *Id*.

[38] *Id*.

[39] *Id*.

[40] *Id*.

[41] *AI*, *supra* note 36.

[42] *Id*.

network.[43] The third step is to "train the algorithm to create the model," or the artificial intelligence system.[44] The last step is to use and improve the model.[45]

[15] Deep learning has shown significantly superior performance in comparison to other traditional machine-learning approaches.[46] Inspired by the latest understanding of human brain behavior, deep learning "utilizes a combination of multi-layer artificial neural networks and data- and compute-intensive training."[47] Deep learning has been so effective that it has even started "to surpass human abilities in many areas, such as image and speech recognition and natural language processing."[48] Deep learning models are able to process vast amounts of data and are often either unsupervised or only semi-supervised.[49] As noted above, both of the main types of artificial intelligence—machine and deep learning—use data to create systems.[50]

---

[43] *Id*.

[44] *Id*.

[45] *Id*.

[46] Rohan Chikorde, *Deep Learning vs Traditional Machine Learning… Which one I should use?*, LINKEDIN (Aug. 25, 2018), https://www.linkedin.com/pulse/deep-learning-vs-traditional-machine-which-one-i-should-chikorde [https://perma.cc/EH47-RKL9].

[47] *AI*, *supra* note 36.

[48] *See id*.

[49] *Id*.

[50] *See* IBM Cloud Education, *Deep Learning*, IBM CLOUD (May 1, 2020), https://www.ibm.com/cloud/learn/deep-learning [https://perma.cc/96KM-LSB6].

[16] Predictably, with the great power of artificial intelligence comes great risk. One of the heightened risks of using artificial intelligence is the risk of using tainted or dirty data to manufacture and train these systems. Because artificial intelligence systems are often unsupervised or only semi-supervised, the "garbage in, garbage out" phenomenon is a real concern.[51] These dirty data points are used to create artificial intelligence systems, which subsequently predict outcomes, manufacture important intelligence, monitor behavior, and more.[52] It becomes a vicious cycle—dirty data creates dirty systems. It is not hard to see why using dirty data to create artificial intelligence systems is problematic. But the problem with using dirty data to create these systems we rely upon runs even deeper and is more troubling than it may first appear. When looking at how artificial intelligence created from dirty data is used in a predictive policing context, the issues become impossible to ignore. Dirty data in artificial intelligence has become increasingly prevalent within predictive policing.[53] To understand how dirty data and artificial intelligence contribute to the issues inherent in predicative policing, it is important to understand what predictive policing is and the problems it creates.

---

[51] *AI*, *supra* note 36; *see* Jason Compton, *Data Quality: The Risks Of Dirty Data And AI*, FORBES (Mar. 27, 2019, 1:21 PM), https://www.forbes.com/sites/intelai/2019/03/27/the-risks-of-dirty-data-and-ai/?sh=597aa8852dc7 [https://perma.cc/ZD8F-39GJ].

[52] *See* Compton, *supra* note 51; Vijay Kanade, *What Is Machine Learning? Definition, Types, Applications, and Trends for 2022*, SPICEWORKS, https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/ [https://perma.cc/NET3-96TH] (Aug. 30, 2022).

[53] Karen Hao, *Police across the US are training crime-predicting AIs on falsified data*, MIT TECH. REV. (Feb. 13, 2019), https://www.technologyreview.com/2019/02/13/137444/predictive-policing-algorithms-ai-crime-dirty-data/ [https://perma.cc/NLD4-BH4P].

## III. PREDICTIVE POLICING

[17] All data sets are somewhat "dirty" because they are filled with errors and mistakes, yet the government at all levels is increasing its reliance on artificial intelligence technology without addressing these mistakes.[54] While there may be some positives to using artificial intelligence, these systems inevitably target marginalized populations and continue to expand the already systemic inequality within our criminal justice system.[55] As Justice Ginsburg said in her dissent in Arizona v. Evans, "[w]idespread reliance on computers to store and convey information generates, along with manifold benefits, new possibilities of error, due to both computer malfunctions and operator mistakes."[56] What would the Supreme Court Justices in 2023 have to say about the chilling racial discrimination stemming from predictive policing based on dirty data? The following section discusses how the integration of dirty data within artificial intelligence leads to problematic and potentially unconstitutional predictive policing. First, it describes what predictive policing is. Then, it turns to algorithms based on location and persons.

### A. What is Predictive Policing?

[18] Predictive policing is a type of predictive tool under the umbrella of predictive analytics and artificial intelligence.[57] Because predictive tools use and abuse dirty data, intentionally or otherwise, they have a great potential for creating long-lasting damage by perpetuating systemic

[54] *See* Valentine, *supra* note 7, at 388–89.

[55] *See id*. at 365.

[56] Arizona v. Evans, 514 U.S. 1, 26 (1995).

[57] *See* Southerland, *supra* note 16, at 497–500.

racism.[58] Although risk assessments have been a part of the criminal justice system for decades, police departments and courts have increasingly turned to artificial intelligence systems to make those risk assessments in the last few years.[59] Due to budget cuts, efficiency has become the focus of policing.[60] Because cities are allocating less of their budgets for police programs,precincts across the nation have turned to algorithms to do the work instead.[61]

[19] Moreover, the increased use of algorithms in the judicial system can be linked to the widespread belief that these systems are more objective than humans.[62] Evidence has suggested that this assumption is far from true.[63] Human prejudices and biases are baked into these tools because the artificial intelligence systems used by police are created using biased police data.[64] Although the racism may be more subtle due to a phenomenon called tech-

[58] *See* Will Douglas Heaven, *Predictive policing algorithms are racist. They need to be dismantled.*, MIT TECH. REV. (July 17, 2020), https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ [https://perma.cc/V844-ZG3B].

[59] *Id.*

[60] *See id.*

[61] *See id.*

[62] *Id.*

[63] *See* Andrew Guthrie Ferguson, *The Police Are Using Computer Algorithms to Tell If You're a Threat*, TIME (Oct. 3, 2017, 11:29 AM), https://time.com/4966125/police-departments-algorithms-chicago/ [https://perma.cc/F5CL-KUAW]; Southerland, *supra* note 16, at 492–94; Heaven, *supra* note 58.

[64] Heaven, *supra* note 58.

washing, it is still apparent.[65] Troubling is the notion that police themselves may even think that these predictive policing systems are less biased and more fair.[66] In fact, the New York City police department admitted that in the wake of the George Floyd protests, it intended to "fight crime differently . . . with less street-stops . . . while better utilizing data, intelligence, and all the technology at [its] disposal . . . . That mean[t] for the NYPD's part, [it would] redouble [its] precision-policing efforts."[67] This acknowledgement shows that while police may understand that there are problems with predictive policing, they may still believe that the answer to these problems lies within artificial intelligence, while failing to address the underlying issues.

[20] However, biases easily bleed into the data.[68] The unfortunate reality is that the discretionary decisions of human police officers tasked with patrolling and investigating suspected crime distorts the ultimate outcome of the data.[69] As humans, our biases are present in almost all the choices we make, and policing is no exception.[70] While it could be argued that having a personal or locational risk score would lead to less racial profiling, the opposite is actually true.[71] Proponents of predictive policing contend that predictive policing provides an alternative to suspicion based upon police

[65] *Id*. (discussing that tech washing is a veneer of objectivity that covers mechanisms that perpetuate inequities in society).

[66] *See id*.

[67] *Id*.

[68] *Id*.

[69] *See* Heaven, *supra* note 58.

[70] *See* Ferguson, *supra* note 63.

[71] *See* Heaven, *supra* note 58.

stereotyping on race, age, or neighborhood.[72] They also speculate that predictive policing could reduce police suspicion for those who are low or no risk individuals.[73] The problem with these arguments is that they do not consider that artificial intelligence is also stereotyping based on age, race, or neighborhood, which is what police officers are then relying on for their police work.[74] While predictive policing serves as a potential avenue for police officers to feel like they are reducing their personal biases by focusing on individuals on the artificially generated heat list, those personal biases will likely be made up for by the dirty and equally biased data being fed into the predictive policing systems.

[21] Equally as troubling is how these biased data points impact how police interact with citizens on the street. Predictive policing influences who police contact and put under surveillance.[75] It also distorts the day-to-day "decisions about the use of force and reasonable [articulable] suspicion."[76] Knowledge based on predictive policing systems colors criminal suspicion and increases police perception of danger, which then results in more frequent and aggressive interactions in specific locations with those that the systems have deemed high risk.[77] To see why this is so problematic, we now turn to the two types of algorithms used in predictive policing: location-based and person-based algorithms.

---

[72] Ferguson, *supra* note 63.

[73] *Id*.

[74] *See id*.

[75] *Id*.

[76] *Id*.

[77] *See* Ferguson, *supra* note 63.

### B. What are location-based algorithms?

[22] Location-based algorithms utilize "links between places, events, and historical crime rates to predict where and when crimes are more likely to happen."[78] These algorithms identify hot spots and police plan patrols of specific locations based on these AI tip-offs.[79] Multiple cities around the U.S. use PredPol, a popular location-based predictive policing system that breaks locations into 500x500 foot blocks to create a "crime weather forecast" that is updated throughout the day.[80] As one might guess, low-income neighborhoods are often targeted through location-based predictive policing systems, which again perpetuates a vicious cycle of systemic racism.[81] While it may seem to make sense that society would want heavier police presence in places the system deems "high-crime neighborhoods," these so-called high crime areas will continuously (and perhaps unfairly) be labeled as such. The data that creates location-based systems and that is frequently used by police officers is often arrest data.[82] However, the systems do not consider that arrest data is not indicative of actual conviction rates.[83]

[23] Even more concerning are the systems that use data where a call to the police has been made.[84] Once the call becomes a data point that will

[78] Heaven, *supra* note 58.

[79] *Id*.

[80] *Id*.

[81] *Id*.

[82] *Id*.

[83] *See* Heaven, *supra* note 58.

[84] *Id*.

justify police going to a specific neighborhood or targeting specific people, the cycle of data-driven technologies legitimizing problematic policing starts.[85] If police are continually directed to specific locations where they can approach suspects for virtually any reason they want, it makes sense that the data would reflect more arrests and issues in these areas. This would then direct police back to those locations, creating an endless cycle. Arrest data is also used to predict potential crimes, but this data does not easily match up with who is actually committing the crimes.[86] Even when it does, there are many other "socioeconomic reasons why [specific] populations and neighborhoods [would] have higher historic crime rates than others."[87]

[24] As noted above, these predictive systems are easily skewed by arrest rates and other data. According to the U.S. Department of Justice, a black person is more than two times as likely to be arrested than a white person.[88] "A [b]lack person is [also] five times as likely to be stopped without just cause as a white person."[89] Although it is prohibited by law to use race as a predictor in intelligence systems, other variables like socioeconomic background, education, and zip code serve as troubling substitutes.[90] So

[85] *See id*.

[86] *See id*.

[87] Heaven, *supra* note 58. S*ee generally* Sarah Childress, *The Problem with "Broken-Windows" Policing*, PBS: FRONTLINE (June 28, 2016), https://www.pbs.org/wgbh/frontline/article/the-problem-with-broken-windows-policing/ [https://perma.cc/9PEA-PP6Q] (discussing why the popular theory of "broken-windows" policing is not actually helpful in policing minority communities but rather perpetuates the cycle).

[88] *Arrest rates by offense and race, 2020 (rates are per 100,000 in age group)*, OFFICE OF JUV. JUST. & DELINQ. PREVENTION (July 8, 2022), https://www.ojjdp.gov/ojstatbb/crime/ucr.asp?table_in=2 [https://perma.cc/352M-N35T].

[89] Heaven, *supra* note 58.

[90] *Id*.

even though race is not explicitly coded into the systems, certain proxies still produce discriminatory outcomes.[91] While there are similar issues within both person-based and location-based predictive policing, each type has distinct concerns.

### C. What are person-based algorithms?

[25] Person-based policing is arguably more problematic than location-based policing, although there are similar root causes.[92] Person-based policing draws on data about people, including their "age, gender, marital status, history of substance abuse, and criminal record, to predict who has a high chance of being involved in future criminal activity."[93] "Person-based . . . policing [was created] in 2009 as an attempt to [use] a public health approach to violence."[94] Similar to epidemiological patterns being used to show which environmental toxins increase health risks, criminal patterns started being used to predict life risks, like gang activity or getting shot.[95]

[26] Although cities across the nation are secretive about how exactly the processes work, there seems to be some sort of general risk evaluation used to predict who is most likely commit crime.[96] The goal of person-based policing is to identify the predictive risk factors for individuals and then try

[91] *See id*.

[92] *See id*.

[93] *Id*.

[94] *See* Ferguson, *supra* note 63.

[95] *Id*.

[96] *See id*.

to remedy the underlying causes creating that risk.[97] To achieve this goal, algorithms are developed for police to prioritize people most at risk by analyzing "past arrests for violent crime, weapons offenses or narcotics; age at the most recent arrest . . . incidents where the individual was a victim of a shooting or assault[,] and the trend line of criminal activity (whether the rate is increasing or decreasing)."[98] The algorithm then analyzes the variables and gives a relative threat score to determine the likelihood of the person either shooting someone or getting shot.[99] This score places that individual on what is commonly known as "the heat list."[100]

[27] In practice, the personalized heat list score will display on computer dashboards so that a police officer can know the alleged risk of the person they are stopping.[101] Those with high scores guide violence-interruption strategies, which in turn influence who the police contact and scrutinize.[102] The score is also used as an indicator of who should be targeted for proactive police intervention.[103] These interventions vary, but can look like home visits by police officers, police surveillance, or being stopped on the street.[104] Because these predictive policing systems are essentially unable

[97] *Id*.

[98] *Id*. (noting that the younger the age, the higher the risk score that was associated with the person).

[99] Ferguson, *supra* note 63.

[100] *Id*.

[101] *Id*.

[102] *Id*.

[103] *Id*.

[104] *See* Ferguson, *supra* note 63.

to be audited (or rather, companies and precincts are *unwilling* to audit them), there is little ability for questioning or controlling the scoring processes of these heat lists.[105]

[28] Police work can be accomplished using predictions of crime, but relying on an algorithm to rank an individual's likelihood to commit crime creates its own risks. Police point to a high percentage of shooting victims being accurately predicted by the heat list as evidence that the algorithms work. However, counterevidence suggests that not only is the targeting overbroad and ineffective, but there are tens of thousands of people labeled as high-scoring who have no history of prior arrest for violent crimes.[106] No matter who you are or what your past looks like, this amount of police scrutiny backed by technology threatens the personal liberties of all citizens.

[29] Dirty data in artificial intelligence has become increasingly prevalent within predictive policing, reinforcing the problematic system within which police work is done through an endless biased data, biased output loop.[107] When predictive policing systems are informed by dirty artificial intelligence, the policies and procedures built from the data cannot be separated from the systemically flawed, racially biased, and inaccurate results being fed into the system in the first place.[108] As a result, these policing practices and policies then shape the environment and practice by which data is created and collected, which leads to an endless cycle of systemic issues within the predictive policing sphere. Whether or not predictive policing is "successful," it raises real constitutional concerns.

---

[105] *See id*.

[106] *Id*.

[107] Richardson, *supra* note 6, at 41.

[108] *See id*. at 195–96.

## IV. CONSTITUTIONAL CONCERNS

[30] Thus far, the discussion has examined two types of predictive policing: location-based and person-based. As it turns out, both types of predictive policing touch vast and monumental bodies of Fourth Amendment law. The first case, *Katz*, and its progeny, defines when a search has been done.[109] The second case, *Terry v. Ohio*, and its progeny, is not about a reasonable expectation of privacy (as is often claimed) but is instead about a seizure and a frisk.[110] To understand the Fourth Amendment implications raised by predictive policing, we must first look at both doctrines. The discussion below begins by looking at a reasonable expectation of privacy—the threshold question for a search in the first place. The discussion then turns to stop-and-frisk practices. Each of these vast bodies of law are implicated by algorithms and the two different types of predictive policing. To understand how they are connected, we must first understand Fourth Amendment law and its principles.

The Fourth Amendment provides that:

[t]he right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.[111]

---

[109] *See* Katz v. United States, 389 U.S. 347, 353 (1967); *see also* United States v. Jones, 565 U.S. 400, 404 (2012) (providing another example of what constitutes as a search and seizure).

[110] *See* Terry v. Ohio, 392 U.S. 1, 9–10 (1968).

[111] *Fourth Amendment*, LEGAL INFO. INST., https://www.law.cornell.edu/constitution/fourth_Amendment [https://perma.cc/R6XV-GLYC].

[31] In short, the Fourth Amendment helps protect individuals from unreasonable police intrusion.[112] Because the Amendment specifies that a person is protected against unreasonable searches and seizures, Fourth Amendment protections are inapplicable if certain police activity is neither a search nor a seizure.[113] Searches under the *Katz*ian protection of information and *Terry* stop-and-frisk context come with a reasonableness standard.[114] The following two subsections of this paper will analyze these two different avenues of potential Fourth Amendment protection within predictive policing.

### A. Protecting Information: Reasonable Expectation of Privacy Even Without a Trespass

[32] While there are many cases dealing with the reasonable expectation of privacy under a search, this paper will discuss five in greater depth: *Katz v. United States* (1967), *United States v. Miller* (1976), *Smith v. Maryland* (1979), *Carpenter v. United States* (2018), and *Herring v. United States* (2009).[115] The cases outlined below will address the reasonable expectation of privacy in a search, the third-party doctrine, and how the Fourth Amendment doctrine applies to technology.

---

[112] Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 336 (2015).

[113] JEROLD H. ISRAEL ET AL., CRIMINAL PROCEDURE AND THE CONSTITUTION 94 (2021 ed. 2021).

[114] *Katz*, 398 U.S. at 360–61 (Harlan, J., concurring); *Terry*, 392 U.S. at 8–9.

[115] While there are other ways to conduct a valid search under the Fourth Amendment, this discussion is limited to the five mentioned in detail. *Katz*, 398 U.S. 347; United States v. Miller, 425 U.S. 435 (1976); Smith v. Maryland, 442 U.S. 735 (1979); Carpenter v. United States, 138 S. Ct. 2206 (2018); Herring v. United States, 555 U.S. 135 (2009).

### 1. A Reasonable Expectation of Privacy: Katz v. United States (1967) and its Progeny

[33] In *Katz v. United States,* the Supreme Court held that to have Fourth Amendment protection, a person must have an objectively reasonable expectation of privacy, and the expectation must be one that society is prepared to recognize as legitimate.[116] The court further held that electronic eavesdropping is governed by the Fourth Amendment.[117] Katz was convicted of transmitting wagering information by phone from state to state in violation of a federal statute.[118] Because the FBI had attached an electronic listening and recording device to the outside of the booth where Katz placed his call, the FBI overheard this call and sought to use the information as evidence against Katz at trial.[119] Katz argued that these recordings were obtained in violation of the Fourth Amendment, and the Supreme Court agreed.[120]

[34] The Court found that "[t]he [g]overnment's activities in electronically listening to and recording the petitioner's words violated the privacy upon which he justifiably relied while using the telephone booth and thus constituted a search and seizure within the meaning of the Fourth

[116] *See Katz*, 398 U.S. at 360–61 (Harlan, J., concurring) ("My understanding of the rule that has emerged from prior decisions is that there is a twofold requirement, first that a person have exhibited an actual (subjective) expectation of privacy and, second, that the expectation be one that society is prepared to recognize as 'reasonable.'"). *But see Katz*, 398 U.S. at 364–65 (Black, J., dissenting) (disagreeing with the majority's interpretation of the Fourth Amendment and instead interpreting it based on the Framers' intent).

[117] *Id*. at 353.

[118] *Id*. at 348.

[119] *Id*.

[120] *See id*. at 359.

Amendment."[121] The Court noted that the surveillance and intrusion of privacy was so narrowly circumscribed that a magistrate could have approved it for a warrant, meaning there would be appropriate safeguards in place against the sole discretion of an officer.[122] Privacy is the key element in *Katz*ian Fourth Amendment protection: if there is a reasonable expectation of a privacy interest in either a search or seizure, then the actions are subject to Fourth Amendment scrutiny.[123]

[35] In *United States v. Miller*, the Supreme Court significantly and definitively limited what is considered a reasonable expectation of privacy when it created what is now known as the "third-party doctrine."[124] This doctrine states that there is generally no reasonable expectation of privacy in regard to information held by a third-party.[125] In *Miller,* the Court found that there was no reasonable expectation of privacy in subpoenaed documents that Miller shared with his bank. Although the Court in *Katz* previously outlined that "a 'search and seizure' become[s] unreasonable when the Government's activities violate the privacy upon which (a person) justifiably relie(s)," it also emphasized that "(w)hat a person knowingly exposes to the public . . . is not a subject of Fourth Amendment protection."[126] Thus, the Court, after examining whether there was a

[121] Katz, 389 U.S. at 353.

[122] *Id*. at 354.

[123] *See* ISRAEL, *supra* note 113, at 94.

[124] *See Miller*, 425 U.S. at 443–44.

[125] *See id*.

[126] *Id*. at 442 (quoting *Katz*, 389 U.S. at 351, 353).

legitimate expectation of privacy in documents conveyed to third-party banks, found there was not.[127]

[36] As the Court explained in *Miller*, the citizen "takes the risk, in revealing his affairs to another, that the information will be conveyed by that person to the Government."[128] The Court noted that it has repeatedly held that "the Fourth Amendment does *not* prohibit the obtaining of information revealed to a third party and conveyed by him to Government authorities."[129] This is the case even if the information was originally revealed on the assumption that it would only be used for a limited purpose and that the confidence originally put into the third-party will not be betrayed.[130]

[37] Just three years after the *Miller* decision, the Court in *Smith v. Maryland* again relied on the third-party doctrine.[131] The Court found that intrinsic in making a phone call is the fact that every number dialed on the phone will be recorded by the phone company.[132] In applying the two prongs of the test described in *Katz*, the Court rejected the claim that there is a legitimate expectation of privacy regarding the phone numbers dialed on his phone.[133] There is no reasonable expectation of privacy to withhold

[127] *Id*.

[128] *Id*. at 443.

[129] *Miller*, 425 U.S. at 443 (emphasis added).

[130] *Id*.

[131] *See Smith,* 442 U.S. at 744.

[132] *Id*. at 742.

[133] *Id*.

the numbers dialed when making a phone call using a landline phone.[134] The Court reasoned that all telephone users realize they must convey numbers to the phone company, so they know they are sharing these numbers with a third-party.[135] The Court further emphasized that the phone company has facilities for recording this information, and that the phone company does in fact record this information for a variety of legitimate business purposes.[136] Although subjective expectations are not scientifically gauged, "it was too [outlandish] to believe that telephone subscribers, under these circumstances, [had] any general expectation that the numbers they dialed [would] remain secret."[137]

[38] Nonetheless, Justice Marshall's dissent in *Smith* compellingly argued that privacy expectations within the meaning of *Katz* should not depend on "the risks an individual can be presumed to accept when imparting information to third parties," but rather should focus on the "risk [s]he is forced to assume in a free and open society."[138] Justice Harlan reinforced this idea when he stated in *U.S. v. White* that "[s]ince it is the task of the law to form and project, as well as mirror and reflect, we should not . . . merely recite . . . risks without examining the desirability of saddling them upon society."[139]

---

[134] *Id*. at 743.

[135] *Id*. at 742–43.

[136] *See Smith*, 442 U.S. at 742 (alluding to billing as a legitimate business purpose by which phone companies record caller information).

[137] *Id*. at 743.

[138] *Id*. at 750 (Marshall, J., dissenting).

[139] United States v. White, 401 U.S. 745, 786 (Harlan, J., dissenting).

[39] Justice Marshall believed that saying the use of pen registers renders telephone numbers dialed to have no reasonable expectation of privacy is "an extensive intrusion that significantly jeopardizes an individuals' sense of security."[140] One cannot use a phone without sharing the numbers, but one also cannot successfully function in society as it currently exists without using a telephone daily.

[40] In line with Justice Marshall's dissent in *Smith*, the Supreme Court has most recently cut back on the third-party doctrine in *Carpenter v. United States,* a case which redefined the Fourth Amendment concept of a reasonable expectation of privacy.[141] The Court previously said that there is no reasonable expectation of privacy in the information shared with a third-party, but *Carpenter* opened the door to limiting that doctrine.[142] *Carpenter* essentially cuts back on the third-party doctrine, although the holding is limited to cell-cite location information ("CSLI") data specifically.[143]

[41] In *Carpenter*, the FBI obtained around 101 location points per day (12,898 location points total over the course of its investigation) that showed Carpenter was near four separate robbery locations.[144] Although the lower court ruled that Carpenter had no reasonable expectation of privacy in information shared with wireless carriers due to the third-party doctrine, the Supreme Court disagreed.[145] "A person does not surrender all Fourth

[140] *See Smith*, 422 U.S. at 751 (Marshall, J., dissenting).

[141] *See generally Carpenter*, 138 S. Ct. at 2223 (creating a right to privacy within third-party doctrine).

[142] *Id*.

[143] *See id*.

[144] *Id*. at 2212–13.

[145] *Id*. at 2223.

Amendment protection by venturing into the public sphere," and individuals *do* "have a reasonable expectation of privacy in the whole of their physical movements."[146] Although the Court in *Miller* and *Smith* applied the third-party doctrine principles strictly, it finally recognized in *Carpenter* that the digital age had changed the trajectory of Fourth Amendment protections.[147] As famously noted in *Katz*, what one seeks to preserve as private, even in an area accessible to the public, may be constitutionally protected. And the Court revived this principle in *Carpenter*.[148]

[42] Finally, in *Herring*, the Court dealt with a case of dirty data – a warrant that had been rescinded but was not updated in the police database system.[149] The system said Herring had an outstanding warrant, which was incorrect.[150] In response to that incorrect computer information, an officer made an arrest and performed a search incident to arrest, with the whole situation coming down to the fact that the officer had relied on a rescinded warrant.[151]

[43] *Herring* turned the conversation from defining a reasonable expectation of privacy that society is willing to accept to the exclusion of evidence that is collected in violation of the Fourth Amendment.[152] The

[146] *Carpenter*, 138 S. Ct. at 2217.

[147] *Id*. at 2220; *see Jones*, 565 U.S. at 417.

[148] *Carpenter*, 138 S. Ct. at 2217 (citing *Katz*, 389 U.S. at 351–352).

[149] *See generally Herring*, 555 U.S. at 135 (explaining that the erroneous arrest warrant led to obtaining evidence in violation of the Fourth Amendment).

[150] *Id*. at 137–38.

[151] *Id*. at 138.

[152] *See id*. at 139.

exclusionary rule says that evidence obtained in violation of the Fourth Amendment is ordinarily inadmissible in a criminal trial.[153] There are strong policy considerations behind the exclusion of evidence based on bad faith actions by the police.[154] These policy concerns are why the doctrine stands, even though the alleged criminal could at times go free simply because the "constable has blundered."[155] However, suppression of evidence "is not an automatic consequence of a Fourth Amendment violation."[156] Deterrence is the key consideration when applying the exclusionary rule, with the question turning on whether the culpability of the police and the potential exclusion of evidence is sufficient to deter wrongful police conduct.[157]

[44] In *Herring*, the majority found that the error in making an arrest based on an outstanding warrant due to negligent bookkeeping "was the result of isolated negligence attenuated from the arrest."[158] At the same time, the Court felt that excluding the evidence would not have a strong enough deterrent effect.[159] The majority in *Herring* ultimately agreed with the lower court, which found that "the conduct in question [wa]s a negligent failure to act, not a deliberate or tactical choice to act."[160] Important to this discussion

[153] Wolf v. Colorado, 338 U.S. 25, 28 (1949).

[154] *See id*.

[155] *See* Mapp v. Ohio, 367 U.S. 643, 659 (1961).

[156] *Herring*, 555 U.S. at 137.

[157] *See id*. at 143.

[158] *Id*. at 137.

[159] *See id*. at 138–39.

[160] United States v. Herring, 492 F.3d 1212, 1218 (11th Cir. 2007).

is *why* the conclusion was appropriate. To the Court, the deterrence benefit of suppressing the evidence did not outweigh the cost, and found that the police made the mistake while acting in good faith.[161]

[45] Notably, however, the Court summed up their opinion by explaining that if there was evidence of "a case where systemic errors were demonstrated, it might be reckless for officers to rely on an unreliable warrant system."[162] It seems to follow that it may be reckless for officers to rely on an unreliable predictive policing system that directly leads to arrests, searches, and seizures. However, the Court failed to acknowledge such possibilities within police databases.

[46] In contrast to the *Herring* majority's unwillingness to address the reliability of police databases, Justice Ginsburg seems to almost have anticipated this issue in her dissent. She was able to see, even then, that "[i]naccuracies in expansive, interconnected collections of electronic information raise grave concerns for individual liberty."[163] As Justice Ginsburg pointed out, the exclusionary rule serves two main purposes: (1) to compel respect for the constitutional guarantee of Fourth Amendment protections in the only effective way possible— "by removing the incentive to disregard it," and (2) to enable the judiciary to assure citizens "that the government w[ill] not [benefit] from its lawless behavior, thus minimizing the risk of seriously undermining… trust in the government."[164] In addition to the two widely accepted purposes of the exclusionary rule, she suggested a new, policy-centered purpose: to monitor and manage the performance of

[161] *See Herring*, 555 U.S. at 138–39; *see also* United States v. Leon, 468 U.S. 897, 922 (1984).

[162] *Herring*, 555 U.S. at 146.

[163] *S*ee *id*. at 155 (Ginsburg, J., dissenting).

[164] *Id*. at 152.

database systems upon which law enforcement rely.[165] This policy-driven approach to Fourth Amendment protections shows exactly why Fourth Amendment protections *do* extend to predictive policing.

### 2. Applying Katz v. United States and its Progeny to Predictive Policing

[47] The cases outlined above have addressed the reasonable expectation of privacy in a search, with some even going as far as addressing how Fourth Amendment doctrine applies to different forms of technology. However, Fourth Amendment protection has never been directly extended by the Supreme Court to predictive policing, despite its obvious applicability. This section will draw out the connections between a reasonable expectation of privacy, as defined throughout history, and predictive policing.

[48] Since 1967, the Supreme Court has addressed the reasonable expectation of privacy in a search in numerous cases, starting with *Katz.* The Supreme Court's holding in *Katz* was extremely important because it meant that a search could occur without a physical intrusion into a constitutionally protected area.[166] By changing the standard instead to an infringement upon a justified expectation of privacy, the scope of the Fourth Amendment was broadened.[167] This is particularly relevant to predictive policing, as the intrusion of privacy from these policing systems goes beyond an intrusion just against one's physical movements. However, within the broader protections of the post-*Katz* Fourth Amendment interpretation, the Supreme Court has still taken a narrow view of what

[165] *See id*. at 153–54.

[166] *See Katz*, 389 U.S. at 353.

[167] *See id*. at 352–53.

justifies a reasonable expectation of privacy.[168] And while the Court has limited the reach of the third-party doctrine defined in *Miller* with its subsequent holding in *Carpenter*, it has only addressed the doctrine further by limiting the third-party doctrine in relation to CSLI data specifically.[169] Focusing on the reach of *Miller* is important because data used for predictive policing purposes may also come from partnerships with third-party data collectors.[170]

[49] As discussed in Part II, data collected for predictive policing includes everything from arrest records to 911 call logs. Much of this information is inevitably shared with third-party data collectors. By definition, does this mean there is absolutely no Fourth Amendment protection because this data was at one point shared with a third-party? It seems inequitable and unfair that the Court would conclude that any and all information that ever makes its way to the police in the wake of big data and artificial intelligence would have no reasonable expectation of privacy. And if the Court did indeed find that none of this data would have a reasonable expectation of privacy, is this the future for policing we want?

[50] A decision like the one contemplated above would seem especially unfair considering that most of the data is distinguishable from that shared with a third-party, like a bank, as was seen in *Miller*.[171] Sharing information with a bank is a purposeful choice; a person, after careful consideration,

[168] *See id*. (holding that a non-physical invasion upon a justified expectation of privacy violates the Fourth Amendment).

[169] *See Carpenter*, 138 S. Ct. at 2223.

[170] *See* WALTER L. PERRY, ET AL., PREDICTIVE POLICING, THE ROLE OF CRIME FORECASTING IN LAW ENFORCEMENT OPERATIONS 84 (2013), https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf [https://perma.cc/7XSM-E4A5 ].

[171] *See Miller*, 425 U.S. at 442.

chooses a bank they feel comfortable working with and proceeds to share confidential information and documents with that bank. Individuals take a conscious risk that the information contained within the documents will be shared. In contrast, data based on past arrests for violent crime, weapons offenses or narcotics, age at the most recent arrest, incidents where the individual was a victim of a shooting or assault, and the trend line of criminal activity is not something a citizen knowingly chooses to share with a third-party.[172]

[51] While this may not have been dispositive under Supreme Court precedent,[173] big data and predictive policing is sufficiently different from the type of third-party information sharing previously considered. The same analysis—that there is no reasonable expectation of privacy for information shared with a third party—should not apply to predictive policing, where data is inevitably shared without consent.

[52] Furthermore, the use of data in predictive policing is also divergent enough from that seen in *Smith* to be constitutionally significant. In *Smith*, the first prong of *Katz* was met because Smith had an objectively reasonable belief that the phone numbers he dialed would remain private, but that expectation was not one that society was willing to recognize as legitimate.[174] The Court instead focused on how the petitioner voluntarily conveyed information to the government by using a phone that provided the phone company with recordable data, and that Smith either knew or should have known that this information could be shared.[175] Data plugged into

[172] Ferguson, *supra* note 63.

[173] *See Smith*, 442 U.S. at 743–44; *Miller*, 425 U.S. at 443 (explaining how Supreme Court case law does not protect information voluntarily given to third parties under the Fourth Amendment).

[174] *See Smith*, 442 U.S. at 743.

[175] *See id*. at 743–44.

artificial intelligence systems like PredPol, on the other hand, is not voluntarily shared in most cases.[176] Moreover, police precincts and cities are incredibly secretive about what kind of information they use and how they use it.[177] Therefore, it would be next to impossible for the Court to assume that citizens know or should know that all of the data being shared with predictive policing systems is being shared with the government to put them under surveillance in the future.

[53] Furthermore, as Justice Marshall persuasively noted in his dissent in *Smith*, "even assuming . . . that individuals 'typically know' that a phone company monitors calls for internal reasons . . ., it does not follow that they expect this information be made available to the public in general or the government in particular."[178] Predictive policing follows the same line of logic: assuming, arguendo, that citizens did know that the police or third-parties keep track of all data about arrests and crime rates (even dirty data), it certainly does not follow that they expect this information to be made available for location-based and person-based predictive policing. As Justice Marshall explained: privacy is not a discrete commodity that is either possessed absolutely or not at all.[179] Those who are arrested, have their age tracked, call the police, live in crime-ridden areas, or have high-risk crime scores (among other similar data) need not assume this information will be used outside the context of normal record keeping and will be released to other persons for other purposes.[180]

---

[176] *See* Heaven, *supra* note 58 (explaining how police artificial intelligence systems use individuals' personal information without their consent).

[177] *Id*.

[178] *Smith*, 442 U.S. at 749 (Marshall, J., dissenting).

[179] *Id*.

[180] *See id*. at 749–50 (explaining why someone does or should know why their information is being shared with third parties).

[54] Predictive policing is a far more extensive intrusion into an individual's sense of security than phone records. To hold otherwise would ignore the inescapable role that dirty data currently plays in artificial intelligence and predictive policing, as well as the Fourth Amendment interests implicated by unchecked police surveillance.[181] Citizens have no choice at all in what kind of collected data goes into these predictive policing systems.[182] The dirtiness of the data creating these systems magnifies just how little of a choice citizens have.[183] Individuals are not knowingly choosing to share skewed and biased data with third-party data collection services for predictive policing, or any other use. Like phone numbers, privacy in living life free from extremely biased and "dirty" predictive policing systems is of value not only to those engaged in criminal activity, but also to those with nothing to hide. As discussed in Section II, many individuals on heat lists have not committed violent crimes but could find themselves under constant surveillance, among other consequences. The prospect of unregulated governmental monitoring under the veil of Fourth Amendment constitutionality should "prove disturbing even to those with nothing illicit to hide."[184] It is important that the Court examines the desirability of saddling the risk of sharing information with the police onto citizens and not merely recite that citizens accept this risk by simply existing in society.[185] It may not have been a consideration in 1979, but it certainly is, and should be, a central concern now.

---

[181] *See id*. at 751.

[182] *See generally* Richardson et al., *supra* note 6 at 22 (stating how police departments have minimal to no oversight over what data they collect and use in their predictive policing systems).

[183] Heaven, *supra* note 58 (stating the problem with predictive policing systems is the type of data being used in the algorithms).

[184] *Smith*, 442 U.S. at 751 (Marshall, J., dissenting).

[185] *See id*. at 750.

[55] Additionally, what the Court found distinguishable in *Carpenter* is exactly what sets the practice of predictive policing apart from what the Court has deemed not to be a Fourth Amendment violation in the past. "Prior to the digital age, law enforcement [could] have pursued a[n] [individual] for a brief stretch [of time] but doing so for any extended period of time was difficult and costly and therefore rarely undertaken."[186] Now, law enforcement can and does secretly monitor and catalogue movements by an individual through CSLI data, GPS, or predictive policing systems like PredPol.[187]

[56] In *Carpenter*, mapping a cell phone's location over the course of 127 days provided an all-encompassing record of the holder's whereabouts that constituted an impermissible search without a warrant.[188] Predictive policing is similarly all-encompassing and would negate any anticipation of privacy in an individual's location or whereabouts.[189] In *Carpenter*, the time-stamped data was said to provide an intimate window into a person's life, with locations holding the "privacies of life" for many Americans.[190] Similarly, predictive policing also "provides an intimate window into a

[186] *Carpenter*, 138 S. Ct. at 2217.

[187] *See Jones*, 565 U.S. at 404 (describing how the government occupies private property for the purposes of obtaining information); *Carpenter*, 138 S. Ct. at 2212 (describing how law enforcement can lawfully obtain cell phone information in criminal investigations).

[188] *Carpenter*, 138 S. Ct. at 2222–23.

[189] *See id*. at 2212. (explaining how police obtain all-encompassing CSLI warrants for preliminary investigatory matters).

[190] *Id*. at 2217.

person's life, revealing not only his particular movements", but often times "familial, political, professional, religious, and sexual associations."[191]

[57] Because predictive policing leads to surveillance and is extensively used by the police, the "privacies of life" that the Court was so concerned about protecting in *Carpenter* are very much at risk with predictive policing. Although predictive policing may not entail the same level of precision as CSLI data, individual liberties and the right to privacy are surely at stake. Like CSLI, predictive police systems are remarkably cheap and efficient compared to traditional investigative tools.[192] As with CSLI, with just the click of a button, the Government can access each person's repository of data at practically no expense.[193] Section II notes why predictive policing started in the first place: to be more efficient and cut costs. As is noted by the majority in *Carpenter,* predictive policing continually logs data on virtually whomever it finds fit, not just those who might justifiably be under investigation. [194] The newfound information capacity runs against everyone.[195] Moreover, whenever a suspect does something that does not rise to the level of probable cause, the police are currently able to call upon the results of predictive policing practices and surveillance without regard

[191] *See id*. (noting how various types of data can provide an intimate window into an individual's life).

[192] *See generally* Heaven, *supra* note 58 (discussing generally the issues with predictive policing and its impacts on minority communities).

[193] *See Carpenter*, 138 S. Ct. at 2217–18.

[194] *Id*. at 2218; Southerland, *supra* note 16, at 505.

[195] *See* Southerland, *supra* note 16, at 505 (describing how different poor data inputs will create bad outputs that may harm communities).

to the constraints of the Fourth Amendment—something the majority explicitly warned against in *Carpenter.*[196]

[58] The Court was sure to include in *Carpenter* that the rule adopted "must take into account of more sophisticated systems that are already in use or in development."[197] It compared CSLI data at the start of the decade as "rapidly approaching that of GPS-level precision."[198] The majority even went so far as to reject the contention of the government and dissent that systems less precise than GPS *should* be allowed.[199] In the same vein, predictive policing systems can narrow down the level of precision to 500x500 feet.[200] Is this not precise enough to deserve Fourth Amendment protection?

[59] The majority rejected the argument that the third-party doctrine governed *Carpenter* because the CSLI data did not constitute business records created and maintained by wireless carriers.[201] So too should the Court find that the third-party doctrine does not apply to predictive policing because it uses data that is not business records and which is shared with third-parties.[202] The Court should therefore find a reasonable expectation of

[196] *See Carpenter*, 138 S. Ct. at 2223 (discussing how although the progress of science is a powerful new tool for law enforcement, this tool risks Government encroachment of privacy rights which the Framers drafted the Fourth Amendment to prevent).

[197] *Id*. at 2210.

[198] *Id*.

[199] *Id*. at 2218–19.

[200] Heaven, *supra* note 58.

[201] *Carpenter*, 138 S. Ct. at 2216–17.

[202] *See id*.

privacy. *Carpenter* accounted for the "seismic shifts in digital technology that made possible the tracking of not only Carpenter's location but also everyone else's, [and] not for a [brief] period, but for years and years."[203] As the Court correctly noted, "there is a world of difference between the limited types of personal information addressed [previously] in *Smith* and *Miller* and the exhaustive chronicle of information casually collected by wireless carriers" today.[204]

[60] Allowing law enforcement to rely on predictive policing systems that are fed dirty data to track the movements of specific citizens in order to essentially surveil those citizens however the police sees fit would be not a straightforward application of the third-party doctrine.[205] Instead, this would be a "significant extension of . . . a distinct category of information."[206] *Carpenter* rejects this extension in its narrow holding, and although it limited the scope of the holding to CSLI data, the holding should rationally be extended to predictive policing as well.[207] This application of the *Carpenter* holding would mean that any searches stemming from predictive policing action based on artificial intelligence would require a warrant supported by probable cause.[208]

[203] *Id*. at 2219.

[204] *Id*. at 2210.

[205] *See id*. at 2219.

[206] *Carpenter*, 138 S. Ct. at 2219.

[207] *See id*. at 2220.

[208] *See id*. at 2221 ("[T]he government must generally obtain a warrant supported by probable cause before acquiring such records.").

[61] As *Carpenter* illustrates, the progress of science should not erode Fourth Amendment protections.[209] The Court is, as it has always been, obligated to ensure that as the invention of subtler and more far-reaching means of privacy invasion have been made available to the Government, those technologies do not effectively destroy Fourth Amendment protections.[210] Predictive policing, like CSLI data, "has afforded law enforcement a powerful new tool to carry out its responsibilities," while at the same time risking encroachment of the kind that the framers of the constitution "drafted the Fourth Amendment to prevent."[211] Thus, just like the Court extended Fourth Amendment protections to the depth, breadth, and inescapable and automatic collection of information of CSLI data (despite it being gathered by a third party,) it should also extend these protections to law enforcement based on predictive policing. These systems are "[no] less deserving of Fourth Amendment protection."[212]

[62] Finally, *Herring* proves exactly why Fourth Amendment protection *should* extend to predictive policing. Relying on the warrant may have been a genuine good faith mistake in *Herring*, as there is no evidence to suggest the officer knew or should have known that the warrant upon which he was relying on was no longer valid.[213] However, relying on artificial intelligence and predictive policing systems that have been repeatedly proven to create

[209] *See generally id*. at 2209–11 (describing how Fourth Amendment protections have been eroded by technology).

[210] *Id*. at 2223.

[211] *See generally Carpenter*, 138 S. Ct. at 2223 (describing how the Founding Fathers envisioned the protections of the Fourth Amendment).

[212] *Id*.

[213] *Herring*, 555 U.S. at 147.

inaccurate and completely biased results based on dirty data is a different situation entirely.

[63] While the officer in *Herring* was genuinely unaware of the issues with the warrant and the computer database,[214] the predictive policing problem has been prevalent for many years throughout the nation and should be hard for police officers to ignore. The argument that the police can rely in good faith on a system that is known to be riddled with harmful flaws is too strained to accept. Moreover, if subjective good faith alone were the test, then Fourth Amendment protections would effectively be read out of the Constitution, and "people would be 'secure in their persons, houses, papers, and effects,' only in the discretion of the police."[215] With the advent of predictive policing comes new constitutional implications, and the Court's current analysis of good faith reliance is no longer sufficient (if it ever was) in protecting citizens from unreasonable and overreaching governmental intrusion.[216]

[64] Furthermore, the error in *Herring* did not rise to the level of "deliberate, reckless, or grossly negligent conduct, or in some circumstances *recurring or systemic negligence*" required by the majority, but predictive policing likely does, or could.[217] Willful blindness or intentional ignorance could certainly be classified as deliberate, and the police conduct at issue has been shown to be both recurring and systemic.[218] Even if the conduct does not rise to the level of being deliberate, it is, at the

[214] *Id*.

[215] *Terry*, 392 U.S. at 22 (quoting Beck v. Ohio, 379 U.S. 89, 97 (1964)).

[216] *Id*.

[217] *See Herring*, 555 U.S. at 144 (emphasis added).

[218] *See generally* Heaven, *supra* note 58 (discussing the continued use of predictive policing by law enforcement despite the algorithmic racial biases in these tools).

very least, reckless. Police conduct based on predictive policing systems recklessly disregards the necessity of policing free from programmed biases and issues, making it necessary to subject police action based on predictive policing systems to the exclusionary rule. Dirty data *is* systemic and recurring, and the Court's analysis in *Herring* comes as close to the issue as any Supreme Court case available.

[65] Though the Court was unaware of the future of predictive policing at the time, the *Herring* Court seemed to address issues at the forefront of predictive policing today. *Herring "*do[es] not suggest that all recordkeeping errors by the police are immune from the exclusionary rule."[219] The majority explicitly stated that "[i]f the police have been shown to be reckless in maintaining a warrant system, or to have knowingly made false entries to lay the groundwork for future false arrests, exclusion would certainly be justified under our cases should such misconduct cause a Fourth Amendment violation."[220] Although predictive policing is not identical to a warrant system, the premise is the same: predictive policing very likely lays the groundwork for future false arrests based on evidence obtained incident to an arrest without proper probable cause. Yet in the end, the majority declined to expand its caselaw to the "unreliability of a number of databases not relevant" to the facts of *Herring*.[221]

[66] Even so, the suggestion of Justice Ginsburg in *Herring* (though not accepted by the majority) applies almost exactly to the problem of predictive policing. She compared the situation to *respondeat superior* liability, explaining that "[j]ust as the risk of *respondeat superior* liability

[219] *Herring,* 555 U.S. at 146.

[220] *Id*. at 146.

[221] *Id*. at 146–47 (citing Arizona v. Evans, 514 U.S. 1, 17 (1995) (O'Connor, J., concurring); Hudson v. Michigan, 547 U.S. 586, 604 (2006) (Kennedy, J., concurring)).

encourages employers to supervise . . . their employees' conduct [more carefully], so the risk of exclusion of evidence encourages policymakers and systems managers to monitor the performance of the systems they install and the personnel employed to operate those systems."[222] Justice Ginsburg perfectly articulated why it would be both logical and equitable to use the exclusionary rule as an incentive to ensure law enforcement properly monitors and manages the predictive policing systems they rely on. As Justice Ginsburg points out, the Court's majority opinion in *Herring "*underestimates the need for a forceful exclusionary rule and the gravity of [the] recordkeeping errors in law enforcement."[223]

[67] The calculus is simple: the exclusionary rule could provide some incremental deterrent: if police were required to exclude evidence obtained solely from predictive policing (which uses data riddled with flaws) then police would be deterred from relying on it. The alleged social cost would be that police may not be able to police as efficiently or have surveillance on whomever they see fit based on the data provided by the systems.[224] While PredPol and other predictive policing systems have been successful at times in preventing crime or stopping the crime before it occurs,[225] the deterrence benefits still seem to outweigh the costs. Therefore, the benefit of deterrence for police relying almost solely on faulty data greatly outweighs the cost of police having to find a fairer, and constitutional, way to police. If police continue to use predictive policing systems, subjecting evidence to the exclusionary rule would at the very least incentivize them to scrub or cleanse the data or use the systems in a better way.

---

[222] *Id*. at 153–54 (quoting *Evans*, 514 U.S. at 29 n.5 (Ginsburg, J., dissenting)).

[223] *Id*. at 150 (Ginsburg, J., dissenting).

[224] *See id*. at 141 (citing United States v. Leon, 468 U.S. 897, 908 (1984)).

[225] Heaven, *supra* note 58.

[68] Consider the following hypothetical.[226] Ben, a man who has previously been convicted of only drug related misdemeanors, and never of a dangerous crime or felony, hears a knock on the door. Ben lives in a high-crime neighborhood in a city that has a high murder rate. Nonetheless, Ben is not involved in the violence surrounding him. Ben opens the door and is surprised to see two police officers, along with two men not in uniform, one of whom is a social worker. However, no one accuses Ben of breaking the law and they are not there to arrest him – they are there to inform him that a predictive policing algorithm had predicted he would be involved in a shooting. Although the officers are unsure whether he will be the shooter or the victim, they are confident based on their data and his proximity to other shootings that he will be involved in one in the future. Police explicitly warn Ben that they will be watching him from here on out, since the algorithm indicates Ben was more likely than 99% of Chicago citizens to either shoot someone or be shot. Ben is confused, because he has no violent history and there is no reason police should be showing up at his door to declare him a potential threat.

[69] Ben notices that ever since Chicago P.D. visited him, they start hanging out where he works, question co-workers about his whereabouts, and look for opportunities to stop him. Ben sees officers continuously hanging around and waiting for him, ready to search and seize him at a moment's notice. One day, the police decide to act. They show up at Ben's workplace (without a warrant) and ask him to open up the company's safe. He doesn't have a key to do so, but once his boss arrives, he opens it. Inside, the officers find a small amount of marijuana and a pipe. Ben is arrested and charged with possession. Because Ben does not have the funds, he does not decide to fight the charge. Police originally approached Ben because they thought he would be involved in a shooting. Now they are charging him

[226] Matt Stroud, *Heat Listed*, THE VERGE (May 24, 2021, 10:00 AM), https://www.theverge.com/c/22444020/chicago-pd-predictive-policing-heat-list [https://perma.cc/YD9S-UXLV] (using Robert McDaniel's story of predictive policing in Austin, Chicago in the following hypothetical).

with another low-level, non-violent offense. However, because Ben was arrested, he will now again be entered into the predictive policing system as a data point. The police will now have an even further reason to continue to surveil Ben, and any of Ben's neighbors, or those in similar circumstances as well.

[70] For too many citizens, the hypothetical above is real.[227] The police are able to target whomever they please for crimes that have not yet occurred, and then find a way to justify their actions after they make an arrest.[228] However, had the police been relying on accurate data, Ben likely would never have found himself on the heat list, because he is a non-violent offender who was not involved in serious crime. This scenario seems all too similar to *Herring*.[229] Herring had a history with the police and on the day of his arrest, he was being followed just like any other day.[230] Herring had not been engaged in any observable criminal behavior at the time of arrest.[231] A flawed database gave the police reason to arrest him, at which point they were able to search him incident to arrest and find a firearm and drugs.[232] In the current hypothetical, the database gave police a reason to approach Ben without a warrant and search his place of work, which then gave them probable cause for arrest. As Justice Ginsburg contended in her

[227] *See id*.

[228] *See id*.

[229] *See generally Herring*, 555 U.S. at 137–38 (holding that Herring's initial arrest was unlawful because the warrant had been rescinded, but due to negligent bookkeeping, the arresting officer had a reasonable belief that the warrant was outstanding).

[230] *See id*. at 137.

[231] *See id*.

[232] *Id*.

dissent in *Herring*, accurate data within law enforcement is of paramount importance, and the deterrence value is strong.[233]

[71] Law enforcement has an increasing supply of information within easy reach. Thus, the risk of error in use of these databases is not a small one.[234] Justice Ginsburg correctly argued that
[i]naccuracies in expansive, interconnected collections of electronic information raise grave concerns for individual liberty. 'The offense to the dignity of the citizen who is arrested, handcuffed, and searched on a public street simply because some bureaucrat has failed to maintain an accurate computer data base' is evocative of the use of general warrants that so outraged the authors of our Bill of Rights.[235]

[72] Regardless of the majority's refusal in *Herring* to address the unreliability of computer databases used by police at the time, the Court can no longer ignore the problem. The Court should address the important questions raised by the unique problems of predictive policing today: how do predictive policing systems reconcile with the values of the Fourth Amendment? Do we want to prioritize law and order over individual liberties? As Justice Frankfurter persuasively noted in *Wolf*, "[t]he security of one's privacy against arbitrary intrusion by the police—which is at the core of the Fourth Amendment—is basic to a free society."[236]

[73] As a society, we should not stand for the practice of unfettered police action against innocent (or guilty) citizens. Allowing the police to

[233] *See id*. at 150–51 (Ginsburg, J., dissenting).

[234] *See Herring*, 555 U.S. at 155 (Ginsburg, J., dissenting).

[235] *Id*. at 155–56.

[236] *Wolf*, 338 U.S. at 27.

circumvent the normal process of investigating crimes that have already happened[237] through reliance on artificial intelligence systems that predict who might commit crimes in the future opens the door to erasing the constitutional right of innocence until proven guilty. Predictive policing provides justification for the presumption that any citizen who finds themselves subject to these policing tactics is guilty until proven innocent. Moreover, it is incorrect to argue, as the government in *Herring* does, that police "have no desire to send officers out on arrests unnecessarily, because [doing so] consume[s] resources and [puts] officers in danger."[238] On the contrary, *Herring* and other data support the proposition that the subjective intent of police is irrelevant. Because the officer wanted to arrest Herring, law enforcement consulted the arrest data system to "legitim[ize] his predisposition."[239] If officers want to arrest someone, are they simply required to check their heat list?

[74] Justice Ginsburg's dissent in *Herring* is strikingly relevant to the issue of predictive policing, and shows that whether it be 2009 or 2023, these issues are real and are becoming more pervasive in society as technology improves. Justice Ginsburg saliently stated that "[n]egligent recordkeeping errors by law enforcement threaten individual liberty, are susceptible to deterrence by the exclusionary rule, and cannot be remedied effectively through other means."[240] Therefore, the dissent in *Herring* shows that the cost *not* to exclude evidence obtained unconstitutionally based on predictive policing is one that society should not be willing to bear. Citizens are entitled to a reasonable expectation of privacy, and in certain

[237] Which brings up its own, distinctly troubling issues.

[238] *See Herring*, 555 U.S. at 156 (Ginsburg, J., dissenting).

[239] *Id*.; *see* Whren v. United States, 517 U.S. 806, 812, 814 (1996).

[240] *Herring*, 555 U.S. at 157 (Ginsburg, J., dissenting).

cases, that expectation extends to information shared with third parties. The third-party doctrine should not erode Fourth Amendment protections, and when it does, evidence obtained unconstitutionally as a result should be subject to the exclusionary rule.[241]

### B. Terry Stop-and-Frisk

[75] The next avenue for potential constitutional protection against predictive policing is under *Terry* stop-and-frisk. Although the standard for a stop-and-frisk is already less than what is required for probable cause, this section argues that predictive policing encroaches upon the reasonable expectation of privacy held during a stop-and-frisk. The Supreme Court has also developed a "reasonable articulable suspicion" standard to stop and seize a person temporarily – this standard was created in *Terry v. Ohio.*[242]

#### 1. Stop-and-Frisk: Terry v. Ohio (1968)

[76] In *Terry,* the Supreme Court created a new (and less stringent) avenue for Fourth Amendment protection.[243] The Supreme Court de-coupled the reasonable expectation of privacy in the search context from the reasonableness of a seizure in the stop-and-frisk context.[244] The Court made this change when it created a reasonable articulable suspicion standard to

[241] *See generally Smith*, 422 U.S. at 735 (discussing how application of the Fourth Amendment depends on whether the person invoking its protection can claim a reasonable expectation of privacy that has been invaded by the government).

[242] *See Terry*, 392 U.S. at 10–11, 19, 21.

[243] *Id*. at 30–31.

[244] *See generally id*. at 30 (holding that a police officer who observes unusual and possible criminal conduct is entitled to conduct a carefully limited search of a person).

justify a brief detention.[245] This standard requires a lower showing than probable cause: the police may not constitutionally stop, seize, or search individuals without having reasonable articulable suspicion that the individual is committing, is about to commit, or has committed a crime.[246] Similar to *Katz*, there is a dual inquiry in deciding if the police acted reasonably: (1) was the officer's action justified at its inception, and (2) was it reasonably related in scope to the circumstances that justified it in the first place?[247] *Terry* stops are recognized as seizures and frisks are recognized as searches, although both are less intrusive than their comparative counterparts.[248]

[77] In simple terms, a stop-and-frisk is when the police are "allowed to 'stop' a person and detain him briefly for questioning upon suspicion that he may be connected with criminal activity."[249] If there is reasonable articulable suspicion that the person may be armed, "the police . . . have the power to 'frisk' him for weapons."[250] "If the 'stop' and 'frisk' give[s] rise to probable cause to believe . . . the suspect has committed a crime, then the police [are] empowered to make a formal 'arrest' and a full incident 'search' of the person."[251]

---

[245] *See id*. at 21–22.

[246] *See id*. at 25–27.

[247] *Terry*, 392 U.S. at 19–20.

[248] *See generally id*. (finding that the distinctions of classical 'stop-and-frisk' theory divert attention from the central inquiry of the reasonableness of the governmental invasion of a citizen's personal security and that a limited search for weapons may be characterized as something less than a 'full' search).

[249] *Id*. at 10.

[250] *Id*.

[251] *Id*.

[78] As noted above, there are two relevant prongs for finding a police stop-and-frisk justifiable: (1) if it was justified at its inception, and (2) whether it was reasonably related in scope to the circumstances which justified the interference in the first place.[252]

[79] For a stop-and-frisk to be justified at its inception under the first prong in *Terry*, a police officer must simply believe that there is suspicion for someone to commit a crime either in the past, present, or future.[253] In sum, the police just have to be able to point to *something*. Moreover, this is an objective standard, meaning the motivation for the stop is inconsequential so long as there is reasonable articulable suspicion.[254] This suspicion can come from the crime, the circumstances around the stop, or a situation that develops once the stop is under way.[255]

[80] Once there is reasonable articulable suspicion to stop the person under the first prong, the second prong is met so long as there is reasonable articulable suspicion that the person is armed and dangerous.[256] Once the second prong is met, the individual may be frisked.[257] Therefore, once both prongs are met and the stop-and-frisk can be articulated as reasonable based

[252] *Terry*, 392 U.S. at 19–20.

[253] *See id*. at 26–27.

[254] *Id*. at 28 (stating that the police officer's hypothesis that the petitioner was contemplating daylight robbery justified the inception of the stop); *Whren*, 517 U.S. at 812 (finding that an officer's motive does not invalidate objectively justifiable behavior for performing a stop).

[255] *Terry*, 392 U.S. at 10–11.

[256] *Id*. at 24.

[257] *Id*.

on police suspicion, the citizen may be constitutionally stopped and frisked.[258] While this is a clear standard by way of legal rule, it does not translate well to peace of mind for people of color and other populations who are fearful of police presence, and for good reason.[259] For those who are already singled out and subjected to law enforcement's biases on a regular basis, it is not reassuring that they may be stopped on the street and subjected to a frisk for anything that is reasonably related in scope to why they were stopped in the first place.[260] Additionally, with predictive policing in the mix, the Court must face that there is now data (albeit inaccurate and flawed) that the police may "reasonably" rely on to justify the initiation of a stop and limited search for whomever they deem suspicious. Although a frisk may only be for weapons and is not as permissive as a full search incident to arrest, it is still a "serious intrusion."[261]

[81] *Terry* stop-and-frisk raises similar concerns about enabling the police to expand their already broad power in an arguably unconstitutional way. As previously noted, *Terry v. Ohio* was an important case, because it is the first exception to the probable cause requirement normally required of a stop and search.[262] *Terry* addresses the "serious questions concerning the role of the Fourth Amendment in the confrontation on the street between the citizen and the policeman investigating suspicious circumstances."[263] This is why *Terry* easily extends to predictive policing issues as well.

[258] *Id*. at 30–31.

[259] *See, e.g.*, Heaven, *supra* note 58.

[260] *Terry*, 392 U.S. at 26.

[261] *Id*.

[262] *See id*. at 24–25.

[263] *Id*. at 4.

[82] In *Terry*, Officer McFadden ("McFadden") was patrolling in downtown Cleveland when two men, Chilton and Terry, attracted his attention.[264] He only had knowledge of what he observed and had never been acquainted with either of the men prior to the afternoon in question.[265] He approached the men and asked their names, and in response to their mumblings, he grabbed Terry and patted down the outside of his clothing, inside of which he found a firearm.[266] The trial court ruled that "it 'would be . . . beyond reasonable comprehension' to find that McFadden had . . . probable cause to arrest the men before he patted them down for weapons."[267] Nonetheless, the trial court declined to exclude the evidence because the frisk was "essential to the proper performance of the officer's investigatory duties, for without it 'the answer to the police officer may be a bullet, and a loaded pistol discovered during a frisk is admissible.'"[268]

[83] Despite being unaware in 1968 what technological advances would be possible in 2023, the Supreme Court in *Terry* raised many of the issues previously discussed in this article.[269] The Court acknowledged that the exclusionary rule "has been recognized as [the] principal mode of discouraging lawless police conduct."[270] However, the Court also

[264] *Id*. at 5.

[265] *Terry*, 392 U.S. at 5–6.

[266] *Id*. at 6–7.

[267] *Id*. at 7–8 (discussing how if there had been probable cause to arrest the men prior to the pat down, this would have been a lawful search incident to arrest).

[268] *Id*. at 8.

[269] *See generally id*. at 38 (discussing issues pertaining to violations of the Fourth Amendment and probable cause).

[270] *Terry*, 392 U.S. at 12.

acknowledged that the "wholesale harassment by certain elements of the police community, of which minority groups, particularly Negros, frequently complain, will not be stopped by the exclusion of any evidence from any criminal trial." [271] Despite these simultaneously true, yet conflicting statements, the Court aimed to lend assurance to the public by reiterating that under its decision in *Terry*, "courts still retain[ed] their traditional responsibility to guard against police conduct which is over-bearing or harassing, or which trenches upon personal security without the objective evidentiary justification which the Constitution requires."[272] The majority said that "[w]hen such conduct is identified, it must be condemned by the judiciary and its fruits must be excluded from evidence in criminal trials[.]"[273] But is that really what the judiciary has committed to since 1968?

[84] The Court seemed to take the easy way out by addressing the narrow question of "whether it is always unreasonable for a policeman to seize a person and subject h[er] to a limited search for weapons unless there is probable cause for an arrest."[274] The Court explicitly declined to address "the scope of a policeman's power when he confronts a citizen *without* probable cause to arrest h[er]."[275] However, this distinction seems to have done little in the way of protecting individual liberties, especially in the wake of predictive policing technology.

[271] *Id*. at 14–15.

[272] *Id*. at 15.

[273] *Id*.

[274] *Id*.

[275] *Id*. at 16 (emphasis added).

### 2. Application of Terry v. Ohio to Predictive Policing

[85] The serious intrusion of stop-and-frisk upheld by *Terry* has certainly only been magnified by predictive policing. Location-based policing tells police where they should patrol in order to fight crime most successfully.[276] Stationing officers based on these systems has already increased police presence in low income and minority neighborhoods.[277] Coupled with the very low standard required of the police currently necessary for a stop and frisk, civil liberties are at risk. As this paper has shown, dirty data has seriously skewed the accuracy of predictive policing.[278] Dirty data gives biased, flawed, and inaccurate data to artificial intelligence systems that are used to inform predictive policing.[279] Thus, police presence is magnified in specific areas, which are oftentimes places where police already have an increased presence.[280] When applied to the predictive policing context, it is clear that the door has been opened too far; law enforcement has now been provided with what they might view as a concrete reason to believe that a person is about to engage in a crime.[281]

[86] No matter if there is probable cause or merely reasonable articulable suspicion (as is required for a stop and frisk), a citizen's personal security is at risk when stopped and searched, either on the street or in their home.[282]

---

[276] Heaven, *supra* note 58.

[277] *See id*.

[278] Valentine, *supra* note 7, at 367.

[279] *Id*. at 368.

[280] *See* Heaven, *supra* note 58.

[281] Southerland, *supra* note 16, at 501.

[282] *See Terry*, 392 U.S. at 19.

The pro-police decision of the Court to create the reasonable articulable suspicion standard made it possible for police to ensure their own safety when approaching citizens on the street, while at the same time creating the possibility for the erosion of citizens' rights under the Fourth Amendment.[283] When viewing *Terry* under a technological advancement lens, it is almost impossible to see how there could be a limit for what might qualify as "reasonable articulable suspicion."

[87] Compare the facts of *Terry* to the aforementioned hypothetical.[284] The police in that situation had no probable cause to arrest Ben of anything. Otherwise, they would have done so in the first place. Instead, the police relied on circumstances that "reasonably" led them to believe that Ben would at some point be involved in a violent crime. If the police had approached Ben on the street instead of his home, they would then have a proper constitutional basis for accosting Ben, restraining his liberty of movement, and addressing questions to him. Whether police know of the specific person they are suspicious of or can articulate suspicion because a predictive policing system has labeled a specific area as problematic, the power of the police has been overextended. Once again, this kind of police overreach should concern the general population as a whole, not just those with something to hide.

[88] The Court in *Terry* recognized how serious, humiliating, and intrusive a frisk is, despite being less so than an arrest.[285] The majority explained that it is simply fantastic to urge that such a procedure performed in public by a policeman while the citizen stands helpless, perhaps facing a wall with his hands raised, is a 'petty indignity.' It is a serious intrusion

[283] *Id*. at 23.

[284] *See* Ferguson, *supra* note 112, at 401.

[285] *Terry*, 392 U.S. at 24–25.

upon the sanctity of the person, which may inflict great indignity and arouse strong resentment, and it is not to be undertaken lightly.[286]

[89] However, despite the majority's warning that this type of intrusion is not merely a "petty indignity" but rather a serious matter,[287] it seems to have left those who are most at risk of this intrusion without an effective remedy. Both prongs of the articulated test can be satisfied rather easily by police as a result of the continuing sophistication of technology.[288] Arguably more troublesome is the issue that, by creating a reasonable articulable suspicion standard for a stop-and-frisk, it is unlikely that evidence found as a result of predictive policing practices will be subject to exclusion under the Fourth Amendment exclusionary rule.

[90] Furthermore, the finding that the exclusionary rule is not an effective deterrent in these types of situations is misplaced.[289] Under the lens of predictive policing, the deterrence factor of not allowing police to stop-and-frisk whomever they please is arguably greater than the risk that any of the people they are interacting with may have a weapon. The majority in *Terry* explained that "in some contexts the rule is ineffective as a deterrent."[290] In the Court's opinion, the diversity of street encounters between law enforcement and citizens exemplifies the low deterrent value of exclusion.[291] In pointing out that street encounters "range from wholly friendly exchanges of pleasantries or mutually useful information to hostile

[286] *Id*. at 16–17.

[287] *Id*. at 17.

[288] *See id*. at 38 (Douglas, J., dissenting).

[289] *Id*. at 13.

[290] *Terry*, 392 U.S. at 13.

[291] *See id*. at 13–14.

confrontations of armed men involving arrests, or injuries, or loss of life[,]" the Court shows just how disconnected it is from society.[292] Just because some interactions may be relatively pleasant, does not mean that law enforcement would not be deterred by the exclusionary rule. As it currently stands, the main goal of police patrolling and street work has been, and remains, stopping crime and apprehending criminals. It would be naïve to assume otherwise.

[91] More importantly, the Court's opinion incorrectly assumes that police officers' street interactions with people of color or members of marginalized communities would be a "wholly friendly exchange[] of pleasantries" or a sharing of "mutually useful information."[293] While tensions between law enforcement and people of color have always been present, one could argue that the Court may have been less fully aware of these issues in 1968. Nonetheless, in 2023 these issues are glaringly obvious, and this line of reasoning can no longer stand, assuming it ever did.

[92] Finally, Justice Douglas' dissent persuasively articulated why giving police arguably greater power than a magistrate is a "long step down the totalitarian path."[294] Even though Justice Douglas considered that the step may have been desirable to keep up with increased crime rates, which is why predictive policing systems were invented and utilized in the first place, he ultimately concluded that a choice of this magnitude should be properly addressed through a constitutional amendment.[295] "Until the Fourth Amendment, which is closely allied with the Fifth, is rewritten,

[292] *Id*. at 13.

[293] *Id*.

[294] *Id*. at 38 (Douglas, J., dissenting).

[295] *Terry*, 392 U.S. at 38 (Douglas, J., dissenting).

the person and the effects of the individual are beyond the reach of all government agencies until there are reasonable grounds to believe (probable cause) that a criminal venture has been launched or is about to be launched."[296]

[93] Justice Douglas was not blind to the societal pressures that have increased police power to give police an upper-hand and said that those powers "have never been greater" than in 1968.[297] However,
if the individual is no longer to be sovereign, if the police can pick him up whenever they do not like the cut of his jib, if they can 'seize' and 'search' him in their discretion, we enter a new regime. The decision to enter it should be made only after a full debate by the people of this country.[298]

[94] Though predictive policing is not entirely identical to the situation in *Terry*, police motivation, also known as adverse incentives, are an important consideration all the same.[299] Some police action is not going to be susceptible to the exclusionary rule. To the extent that the police have other motives and do not violate civil liberties with the goal of obtaining evidence for prosecution, the Fourth Amendment is not an effective solution.[300] Regardless of how effective the exclusionary rule may be, it is "powerless to deter invasions of constitutionally guaranteed rights where the police either have no interest in prosecuting or are willing to forgo

[296] *Id*. at 39.

[297] *Id*.

[298] *Id*.

[299] *See generally id*. (highlighting the similarities between *Terry v*. *Ohio* and predictive policing).

[300] *Terry*, 392 U.S. at 14–15.

successful prosecution in the interest of serving some other goal."[301] Police use predictive policing not just as a predictive measure, but also as a preventative measure.[302] Thus, this paper concedes that the police use predictive policing for reasons other than successful prosecution.[303]

[95] However, this is not a problem unique to predictive policing. Every issue analyzed under Fourth Amendment doctrine falls victim to this predicament; the exclusionary rule is only a deterrent if there is evidence to exclude from trial in the first place.[304] Nonetheless, to the extent that we have Fourth Amendment protection at all, these constitutional avenues are the way to enforcement, and this paper shows why there is room for that discussion.

## V. CONCLUSION

[96] This paper has argued that artificial intelligence as used in predictive policing implicates Fourth Amendment concerns and has shown that there is indeed room in Fourth Amendment doctrine to regulate dirty data in predictive policing through the exclusionary rule. Nevertheless, if the police do not care about the admissibility of evidence, there is quite little the exclusionary rule or Fourth Amendment protections can do. While the focus of this paper has been the constitutional concerns related to predictive policing, some questions remain. Is predictive policing fair? And are algorithm-driven actions the future society wants for policing?

---

[301] *Id*. at 14.

[302] *See* Valentine, *supra* note 7, at 364.

[303] *See id*.

[304] *See Terry*, 392 U.S. at 12, 14.

[97] The best answer is no, predictive policing is clearly not fair and this is not the future society should want. As noted above, there are extreme issues with how predictive policing disproportionately and unfairly affects people of color.[305] Justice Ginsburg was able to see, even in 2009, that inaccuracies in electronic databases raise grave concerns for individual liberty.[306] Furthermore, as a policy matter, it is bad business for both private citizens and police departments alike to rely on predictive policing systems that are fed dirty data. Relying on dirty data is a harmful practice for citizens, states, and communities; society should collectively strive to uphold the utmost constitutional protections for ourselves and others. This is true even from the perspective of law enforcement — police departments nationwide are losing the trust of citizens and lawmakers and are finding themselves in the middle of a cultural shift.[307] Predictive policing should concern those who are guilty, those who are innocent, and our society as a whole. Everyone loses. Predictive policing funded by dirty data is not supported by the values society should want to be reflected in our criminal justice system.

[98] Recognizing these truths, one city has acknowledged how unfair and problematic predictive policing is and has decided not to reward police work motivated by adverse incentives. Santa Cruz has banned predictive policing altogether, and other states should too.[308] Although Santa Cruz was one of

[305] *See* Valentine, *supra* note 7, at 364.

[306] *See Herring*, 555 U.S. at 155 (Ginsburg, J., dissenting).

[307] *See* Aimee Ortiz, *Confidence in Police Is at Record Low, Gallup Survey Finds*, N.Y. TIMES (Aug. 12, 2020), https://www.nytimes.com/2020/08/12/us/gallup-poll-police.html [https://perma.cc/ET3Q-5S3N].

[308] Kristi Sturgill, *Santa Cruz becomes the first U.S. city to ban predictive policing*, L.A. TIMES (June 26, 2020, 12:19 PM), https://www.latimes.com/california/story/2020-06-26/santa-cruz-becomes-first-u-s-city-to-ban-predictive-policing#_=_. [https://perma.cc/346W-7UZJ].

the first cities to adopt a predictive policing model in the first place, they also became the first city in the country to permanently ban those same systems in 2020.[309] This is encouraging news: if one city can recognize the damaging effects of predictive policing,[310] other cities should certainly be able to follow suit.

[99] The Fourth Amendment is crucial in providing protection against overly intrusive police interactions, and predictive policing should not come with the cost of compromising civil liberties. No matter how we do it, this is a problem worth fixing. Dirty data leads to dirty policing, and we as a society have a duty to clean it up.

[309] *Id*.

[310] *See id*.